Identifying Weekly Trajectories of Pain Severity Using Daily Data from a Mobile Health Study: A Cluster Analysis

Claire L Little, David M Schultz, Thomas House, William G Dixon, John McBeth

Submitted to: JMIR mHealth and uHealth
on: April 28, 2023

Table of Contents

Identifying Weekly Trajectories of Pain Severity Using Daily Data from a Mobile Health Study: A Cluster Analysis

Claire L Little[1] MA, BSc; David M Schultz[2, 3] PhD; Thomas House[4] PhD; William G Dixon[1, 5] PhD; John McBeth[1, 5] PhD

[1]Centre for Epidemiology Versus Arthritis University of Manchester Manchester GB
[2]Department of Earth and Environmental Sciences Centre for Atmospheric Science University of Manchester Manchester GB
[3]Centre for Crisis Studies and Mitigation University of Manchester Manchester GB
[4]Department of Mathematics University of Manchester Manchester GB
[5]NIHR Manchester Musculoskeletal Biomedical Research Unit Central Manchester University Hospitals NHS Foundation Trust Manchester GB

Corresponding Author:
Claire L Little MA, BSc
Centre for Epidemiology Versus Arthritis
University of Manchester
Oxford Road
Manchester
GB

Abstract

Background: People with chronic pain have highlighted a need to forecast variability in their pain severity. We propose a forecasting model for both short-term variability (e.g. daily fluctuations) and longer-term variability (e.g. weekly patterns). For development of this model, clusters of weekly trajectories of pain severity are required, so that future work can predict between-cluster movement and within-cluster variability in pain severity.

Objective: This paper aims to understand clusters of common weekly patterns as a first stage in the development of a pain-forecasting model.

Methods: Data from a population-based mobile health (mHealth) study were used to compile weekly pain trajectories (n = 21,919) and clustered using a k-medoids algorithm. Sensitivity analyses tested the impact of assumptions related to the ordinal and longitudinal structure of the data. The characteristics of people within clusters were examined and a transition analysis was conducted to understand the movement of people between consecutive weekly clusters.

Results: Four clusters of weekly pain severity were identified representing trajectories of no or low pain (n = 1714), mild pain (n = 8246), moderate pain (n = 8376), and severe pain (n = 3583). Sensitivity analyses indicated a four-cluster solution remained suitable under changing assumptions, and resulting clusters were similar to the main analysis, with at least 85% of trajectories belonging to the same cluster as the main analysis. Men spent longer (7.9% of weeks) in the “no or low pain” cluster than women (6.5% of weeks). Younger people (17–24 year olds) spent longer (28.3% of weeks) in the “severe pain” cluster than those aged 65–86 years (9.8% of weeks). People with fibromyalgia (31.5% of weeks) and neuropathic pain (31.1% of weeks) spent longer in the “severe pain” cluster than other conditions, and people with rheumatoid arthritis spent longer (7.8% of weeks) in the “no or low pain” cluster than other conditions. There were 12,267 pairs of consecutive weeks which contributed to the transition analysis. The empirical percentage remaining in the same cluster across consecutive weeks was 66%. When movement between clusters occurred, the highest percentage of movement was to an adjacent cluster.

Conclusions: The clusters of pain severity identified in this study provide a parsimonious description of the weekly experiences of people with chronic pain. These clusters could be used for future study of between-cluster movement and within-cluster variability, in order to develop accurate and stakeholder-informed pain forecasting tools.

(JMIR Preprints 28/04/2023:48582)
DOI: https://doi.org/10.2196/preprints.48582

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Identifying Weekly Trajectories of Pain Severity Using Daily Data from a Mobile Health Study: A Cluster Analysis

Paper Type: Original Paper

Authors:
Claire L Little, MA
ORCID: 0000-0003-2602-0491
Centre for Epidemiology Versus Arthritis,
University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom

Prof David M Schultz, PhD
ORCID: 0000-0003-1558-6975
Department of Earth and Environmental Sciences,
Centre for Atmospheric Science,
University of Manchester, Manchester, United Kingdom
Centre for Crisis Studies and Mitigation
University of Manchester, Manchester, United Kingdom

Prof Thomas House, PhD
ORCID: 0000-0001-5835-8062
Department of Mathematics
University of Manchester, Manchester, United Kingdom

Prof William G Dixon, MRCP, PhD
ORCID: 0000-0001-5881-4857
Centre for Epidemiology Versus Arthritis,
University of Manchester, Manchester, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Prof John McBeth, PhD
ORCID: 0000-0001-7047-2183
Centre for Epidemiology Versus Arthritis,
University of Manchester, Manchester, United Kingdom
NIHR Manchester Musculoskeletal Biomedical Research Unit,
Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Corresponding Author: Claire L Little, claire.little@postgrad.manchester.ac.uk

Abstract

Background: People with chronic pain experience variability in their trajectories of pain severity. Previous studies have explored pain trajectories by clustering sparse data but to understand daily pain variability there is a need to identify clusters of weekly trajectories using daily pain data. Between-week variability can then be explored by quantifying the week-to-week movement between these clusters.

People with chronic pain have highlighted a need to forecast pain variability to reduce pain-related uncertainty. We propose that future work can use clusters of pain severity in a forecasting model for both short-term variability (e.g., daily fluctuations) and longer-term variability (e.g., weekly patterns). Specifically, future work can use clusters of weekly trajectories to predict between-cluster movement and within-cluster variability in pain severity.

Objective: This paper aims to understand clusters of common weekly patterns as a first stage in the development of a pain-forecasting model.

Methods: Data from a population-based mobile health (mHealth) study were used to compile weekly pain trajectories (n = 21,919) and clustered using a k-medoids algorithm. Sensitivity analyses tested the impact of assumptions related to the ordinal and longitudinal structure of the data. The characteristics of people within clusters were examined and a transition analysis was conducted to understand the movement of people between consecutive weekly clusters.

Results: Four clusters of weekly pain severity were identified representing trajectories of no or low pain (n = 1714), mild pain (n = 8246), moderate pain (n = 8376), and severe pain (n = 3583). Sensitivity analyses indicated a four-cluster solution remained suitable under changing assumptions, and resulting clusters were similar to the main analysis, with at least 85% of trajectories belonging to the same cluster as the main analysis. Men spent longer (7.9% of weeks) in the “no or low pain” cluster than women (6.5% of weeks). Younger people (17–24 year olds) spent longer (28.3% of

weeks) in the “severe pain” cluster than those aged 65–86 years (9.8% of weeks). People with fibromyalgia (31.5% of weeks) and neuropathic pain (31.1% of weeks) spent longer in the “severe pain” cluster than other conditions, and people with rheumatoid arthritis spent longer (7.8% of weeks) in the “no or low pain” cluster than other conditions. There were 12,267 pairs of consecutive weeks which contributed to the transition analysis. The empirical percentage remaining in the same cluster across consecutive weeks was 66%. When movement between clusters occurred, the highest percentage of movement was to an adjacent cluster.

Conclusions: The clusters of pain severity identified in this study provide a parsimonious description of the weekly experiences of people with chronic pain. These clusters could be used for future study of between-cluster movement and within-cluster variability, in order to develop accurate and stakeholder-informed pain forecasting tools.

Keywords: mHealth; Pain; Cluster; Trajectory; *k*-medoids; Transition; Forecast

Introduction

Chronic pain (i.e. pain lasting three or more months) is a common symptom of many long-term health conditions [1, 2] and is associated with poor quality of life, poor health outcomes, and low participation in work and social activities [3, 4]. There is substantial day-to-day variability in the severity of pain experienced [5, 6] and people with chronic pain report that this variability leads to feelings of frustration and uncertainty about future pain [7, 8]. Studies have identified associations between pain variability and response to treatment [9] and lower quality of life [10, 11]. However, pain variability remains underestimated by researchers [12].

One way to explore pain variability is to cluster common pain trajectories and to quantify movement between clusters. Previous studies have identified patterns of pain severity by clustering pain trajectories among individuals with chronic pain. These studies have often used sparse data on pain severity collected once per week (e.g. [13]), once per month (e.g. [14]) or less frequently (e.g. [15]). These studies inform our understanding of longer-term experiences of chronic pain, but not the day-to-day experience of pain severity that is important to patients. There is a need to extend this knowledge of pain clusters to within-week pain trajectories. Recent advances in mobile health (mHealth) methods that support the collection of data in the people's own environments [16, 17], often using their own devices such as smartphones and tablets [18], offer the opportunity to capture daily pain severity data.

It is also possible to explore movement between clusters of pain data. For example, Rahman et al. [19] used changes between pain severity scores (not necessarily day-to-day changes) to identify two clusters of low pain volatility and high pain volatility. They then predicted movement between these clusters at six-month intervals. However, there is a need to explore movement between clusters on a shorter time frame.

Once identified, weekly pain trajectories could be forecast. People living with chronic pain have

reported that a pain forecast would reduce unpredictability and could be used to support planning daily activities, such as shopping, chores and social participation [20, 21]. In a research prioritisation study, 75% of respondents to a survey said they would use a pain forecast, and prioritised a model that predicted daily fluctuations (i.e. relatively short-term variability) and pain flares (patterns across multiple days).

We propose three stages to develop a pain forecast (Figure 1). Stage one identifies common weekly trajectories of pain severity using cluster analyses. Stage two investigates day-to-day variability in trajectories of pain severity for individuals within each cluster. These first two stages provide a better understanding of an individual's pain experiences. Stage 3 predicts for an individual their movement between clusters of pain severity across consecutive weeks and future within-cluster day-to-day variability. This study focusses on the first of these stages: clustering trajectories of pain severity. Understanding clusters of weekly trajectories is an important stage in this forecasting model to identify group-level associations that may be masked by population-level analysis.

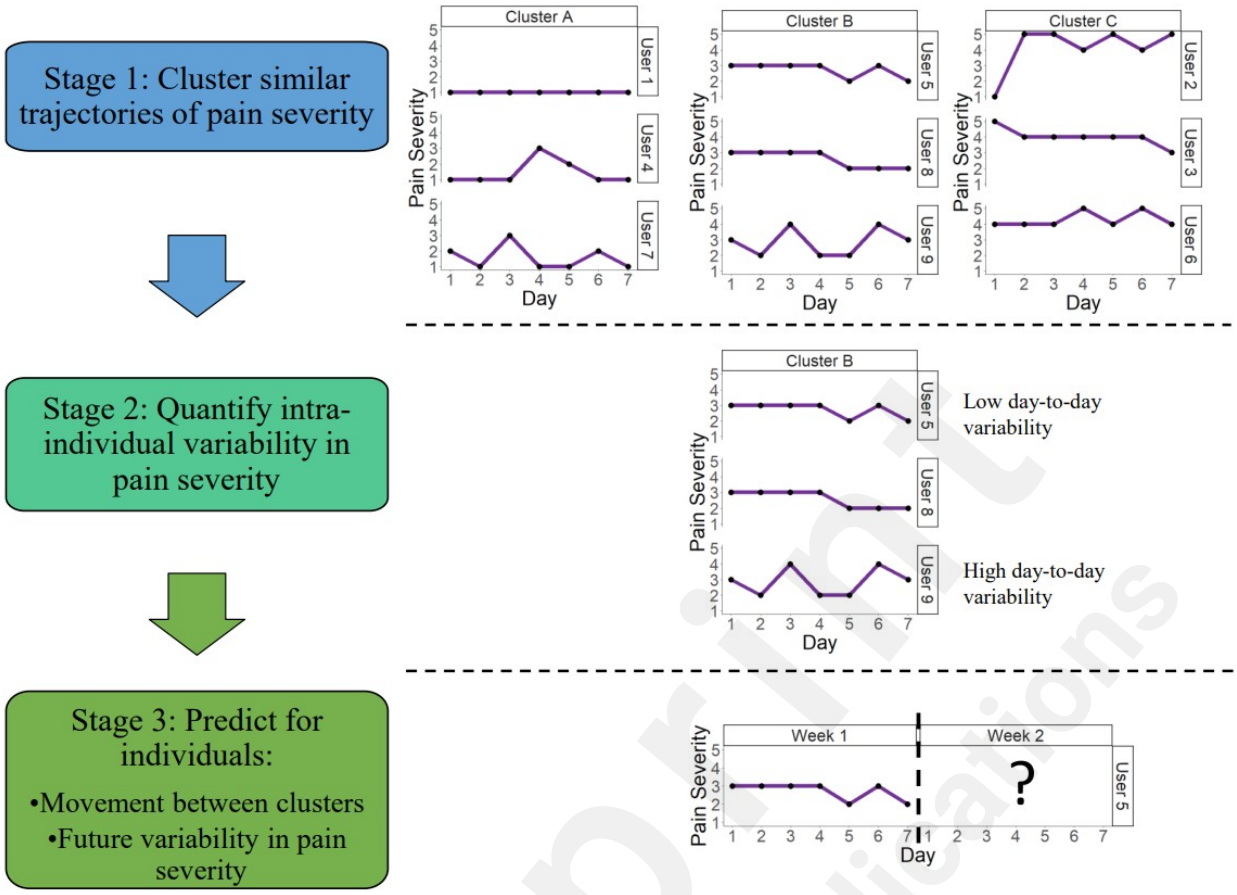


Figure 1: Three stages to build a pain forecast using data from an mHealth study. Data used in this figure are for illustrative purposes only, to provide one example of how data may be used in the pipeline of developing a pain forecast. First, data are clustered to identify common trajectories of weekly pain severity. Second, the remaining variability is explored for each trajectory within a cluster. The process is repeated for each cluster. Third, movement between clusters on consecutive weeks and the amount of day-to-day variability are predicted for an individual. The process is repeated for each individual.

Once collected, there are several challenges to overcome in clustering daily pain severity data for use in a pain-forecasting model. First, patient-generated health data are often collected on an ordinal scale. However, equal intervals between responses cannot be assumed, and using metric models to analyse ordinal data can lead to erroneous errors [22]. Second, data collected are longitudinal, and algorithms used for clustering should respect this longitudinal feature of the data. Third, clusters of pain severity that will be used in a pain-forecasting model should be interpretable to end-users. To

address these challenges requires that we identify and use a suitable method for clustering patient-generated health data. Any assumptions made about the data should be tested in sensitivity analyses to ensure robustness. Observing substantial movement between clusters would suggest the feasibility of forecasting cluster movement in future work. Therefore, understanding the characteristics of individuals who contribute to different clusters and how those individuals move between clusters over time will aid end-user interpretability.

The aim of this study was to understand pain severity clusters in people living with chronic pain. The specific objectives were to (1) use a suitable algorithm to identify the optimum number of clusters of pain trajectories, (2) conduct sensitivity analyses to test assumptions made when clustering data, (3) examine the characteristics of people within clusters, and (4) describe the movement of people between different clusters over time.

Methods

Data source

This study is a secondary analysis of a population-based mHealth study, Cloudy with a Chance of Pain [16, 23, 24]. Study participants were recruited between January 2016 and January 2017 following advertisements on television, radio, and social media. Data collection ended in April 2017, with participants able to contribute data for between 0–15 months. Inclusion criteria required participants to have chronic pain, be aged ≥ 17 years, live in the UK and own an Android or iOS smartphone. Participants downloaded a co-designed mobile phone application (app), provided electronic consent, and provided demographic information including their sex (male or female), year of birth (entered as free text) and pain condition(s) (selected from a list of pre-defined responses e.g., rheumatoid arthritis, fibromyalgia). Daily reports of ten variables were collected, including pain severity. Participants were asked: “How severe was your pain today?” and responded by selecting “no pain” (score 1), “mild pain” (2), “moderate pain” (3), “severe pain” (4) or “very severe pain” (5).

Daily reports of other variables included mood, fatigue, and physical activity, but those were not included in this secondary analysis. Data were collected locally on the smartphone, transferred to an external server where they were anonymised, and then returned to the researchers in anonymised form. Daily data could be contributed for a maximum of 15 months, with participants requested to track symptoms for six months. In total, 10,584 people downloaded the app and recorded their demographic information and at least one record of pain severity. These participants were 81% female (8554 participants), with a mean age of 51 years (standard deviation (s.d.) 12.5 years). On average, these participants contributed pain severity data on 76 days (90% of participants contributed data between 1 and359 days). Previous analysis of this data classified participants as highly engaged (13.6%, 865/6370), moderately engaged (21.8%, 1384/6370), low engaged (39.4%) or tourists (25.4%, 1618/6370) [24].

Ethical considerations

Ethical approval for *Cloudy with a Chance of Pain* was obtained from the University of Manchester Research Ethics Committee (ref: ethics/15522) and from the NHS IRAS (ref: 23/NW/0716). Participants were required to provide electronic consent for study inclusion. Anonymised data were received by the research team. Further ethical approval was not required for the secondary analysis described in the present study.

Data preparation

For this study, weekly trajectories of pain-severity data were used. Trajectories beginning on a Monday were identified, to align data across multiple respondents. This alignment introduced a structure to the data based on the work week, to mitigate the impact of individuals entering the study at different times, and to deal with day-of-the-week effects. A complete participant-week was defined as complete pain-severity data contributed by a single participant during a single calendar week (Monday–Sunday) (Figure 2). Pain-severity data from a complete participant-week were included in

the analysis if (a) the participant had joined the study on, or before, the Monday, (b) the participant had remained in the study on, or after, the following Sunday and (c) the participant had provided complete pain-severity data (i.e. one pain severity score on each of the seven days). Multiple complete participant-weeks could be included in the analysis for each participant (up to 64 weeks due to the length of the study).

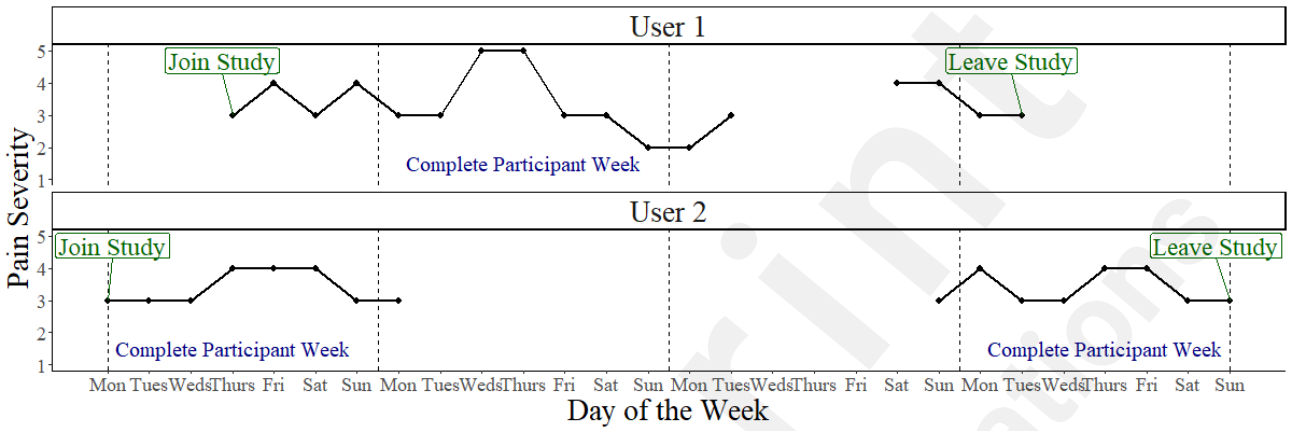


Figure 2: Example selections of complete trajectory-weeks for two participants. The participants join and leave the study at different times. One complete participant-week from User 1 is included in the analysis. Two complete participant-weeks from User 2 are included in the analysis.

Statistical methods

Identifying the optimal number of clusters

Previous studies have used a range of methods to cluster pain severity, including *k*-means clustering [25, 26], hierarchical cluster analysis [27], growth mixture modelling [28–31], latent class growth analysis [13–15, 32, 33], multilevel latent class analysis [34], and group-based trajectory analysis [35–40]. Different approaches have different strengths and kinds of assumption; for example, some may assume that clusters are internally homogeneous or the data overall follow a particular (e.g., linear) form, are continuous or similar. In fact, clustering of ordinal and longitudinal data using a model that explicitly represents these features in a computationally inexpensive way remains a major unsolved methodological challenge. In this study, we choose a method that does not make strong

assumptions about the form or generating mechanism of the data, or the within-cluster variance, and maintains the assumption that the data are ordinal nature while having excellent computational performance and convergence properties. However, the assumption about the longitudinal nature of the data is relaxed.

To identify the optimal number of clusters, data were summarised in feature vectors, compared using the Manhattan (ℓ_1) distance measure, and clustered using an adaptation of the *k*-medoids algorithm, detailed below. The feature vectors were 7-dimensional with entries representing the pain-severity data on each of the seven days in a complete participant week. Using the data directly in this way ensured that feature vectors remained interpretable. The differences between feature vectors were found by calculating the Manhattan distance through entry-wise summation of absolute differences to respect the ordinal nature of the outcome variable. The implementation of the *k*-medoids algorithm used to cluster the feature vectors can be derived as follows. A *k*-medoids algorithm randomly assigns a user-defined '*k*' feature vectors to be the cluster centres (or medoids) and then iteratively (1) assigns each feature vector to the closest medoid and (2) recalculates the medoid of the clusters. The term 'medoid' refers to the use of actual data points as the centres for the clusters [41]. Such use of observed data as centres for the clusters prevents outputs such as "pain severity of 3.2", that might arise if means are used and that are uninterpretable and erroneously assume an interval scale. To implement the *k*-medoids algorithm, the Clustering Large Applications (CLARA) program was used, which was specifically designed for use of large data sets to reduce overall computation time [42, 43].

A *k*-medoids algorithm requires a user-defined value for the number of clusters (*k*) in the data [41]. The implementation of CLARA was therefore repeated for values of *k* from 1 to 20. The output of the algorithm can be sensitive to the random feature vectors selected as the medoids in the first stage of the algorithm. The algorithm was therefore repeated 20 times, once for each value of *k*. At each iteration, the remaining variance within each cluster was calculated as the 'within-cluster sum of

squares' (WSS). The WSS calculates the total remaining distance between pairs of feature vectors in the same cluster. For each value of *k*, the iteration which returns the smallest value of WSS is selected and reported on a plot.

The optimal number of clusters was then selected by three criteria. First, from the plot of *k* against WSS, the optimal number of clusters was chosen visually using the elbow method [44]. While ideally a formal trade-off would be made between model complexity and goodness of fit, there is no clear method to use. Existing methods (e.g. information criteria, silhouette method, gap statistic) can suggest different numbers of clusters [45], possibly due to under-penalising the complexity for datasets of the size used in this study. Therefore, the less formal elbow method allows us to be more explicit in the judgements we make, to resolve the absence of an unambiguous method for learning cluster numbers from data. Second clusters were required to contain 5% of the trajectories, similarly to previous studies [13, 36, 46–49]. Third, cluster solutions were examined for clinical interpretability. For this measure, candidate solutions were examined to ensure meaningful differences between the cluster medoids. Further, the distribution of demographic data of participants contributing trajectories to each cluster were examined to ensure that results reflected expected distributions.

Sensitivity analyses

Three sensitivity analyses were conducted to test assumptions made in the main analysis, with the methodology behind choosing these being to modify assumptions made about the data and to see if broad conclusions were robust. Robust conclusions would be indicative of a strong model-independent signal in the data, even if modified assumptions led to a less interpretable output. The main analysis assumed that data were on an ordered scale but relaxed the assumption that the data were longitudinal.

The first sensitivity analysis maintained the longitudinal nature of the data but implicitly assumed that the outcome variable was on a continuous scale. Feature vectors were compared using the

Euclidean distance, which erroneously assumes regular intervals between values on the pain severity scale. However, the use of Euclidean distance permits the use of the *KmL* package, which specifically clusters longitudinal data [50]. The *KmL* package is an adaptation of the *k*-means algorithm. The *k*-means algorithm is similar to *k*-medoids, but the centre of each cluster is calculated using the mean of the feature vectors assigned to the cluster. The use of the *KmL* package, instead of the CLARA program, and the resulting use of mean trajectories rather than medoid trajectories, were the only adaptions to this sensitivity analysis. The feature vectors, the 20 repetitions of the algorithm for each value of *k*, and the use of elbow method to select *k* remained unchanged.

The second sensitivity analysis relaxed assumptions about both the longitudinal and ordinal nature of the outcome variable. In this sensitivity analysis, the data were not assumed to be longitudinal, and the outcome variable was assumed to be unordered categorical data. A different feature vector was used that converted ordinal pain-severity values into dummy variables using one-hot encoding. In this encoding, there were 35 binary categories, each representing a unique day and pain-severity category. The feature was recorded as '1' if the pain-severity score was seen on that day, and '0' otherwise. In this way, seven of the features were '1' for each complete participant-week. The feature vectors were compared using the *Jaccard distance* typically used for such vectors of binary data. The cluster analysis was then conducted using the CLARA program, in the same manner as described in the main analysis.

The third sensitivity analysis challenged the definition of a Monday–Sunday week when defining complete participant-weeks. Instead, the following analysis was conducted for each of the days (*D*) in the week. Complete participant-weeks were selected from the original data for each participant when there were pain severity data for each day in the *D* to *D*+6 week (e.g. Wednesday to Tuesday week). On each new dataset corresponding to a different *D*, clustering was conducted using the CLARA program at each value of *k* between 1 and 20, as described in the main analysis. Due to the adapted complete participant-weeks, individuals may have contributed different numbers of weeks to

the sensitivity analysis.

For each sensitivity analysis, the optimal number of clusters was calculated. Similar numbers and descriptions of the clusters would provide evidence that the conclusions from the main analysis are robust. Furthermore, for each cluster in the main analysis, the proportion of trajectories allocated to the same cluster in each sensitivity analysis was calculated. A high proportion would further suggest that the results are robust to the assumptions made by using the Manhattan distance and *k*-medoids algorithm in the main analysis.

Description of clusters

Information about the trajectories assigned to each of the clusters in the optimal solution were summarised. First, the number of trajectories assigned to each cluster was reported. Secondly, the clusters were visualised with a spaghetti plot of individual trajectories and the medoid of each cluster. Finally, the average proportion of time spent in each cluster by each participant was calculated. This information was summarised by calculating the mean proportion of time spent in each cluster across demographics (i.e. age, sex, chronic-pain condition or conditions).

Transition between clusters

For the optimal solution of clusters in the main analysis, transition of individuals between clusters on consecutive weeks was examined. To do this, a subset of the total data was used. Complete participant-weeks (“This week”) were retained if the participant had also contributed a complete participant-week in the directly preceding week (“Last week”). A trajectory could be labelled as both “This week” and “Last week” if there were both preceding and succeeding weeks for the individual (Figure 3). The demographic data of participants included in this transition analysis were compared to those included in the main analysis.

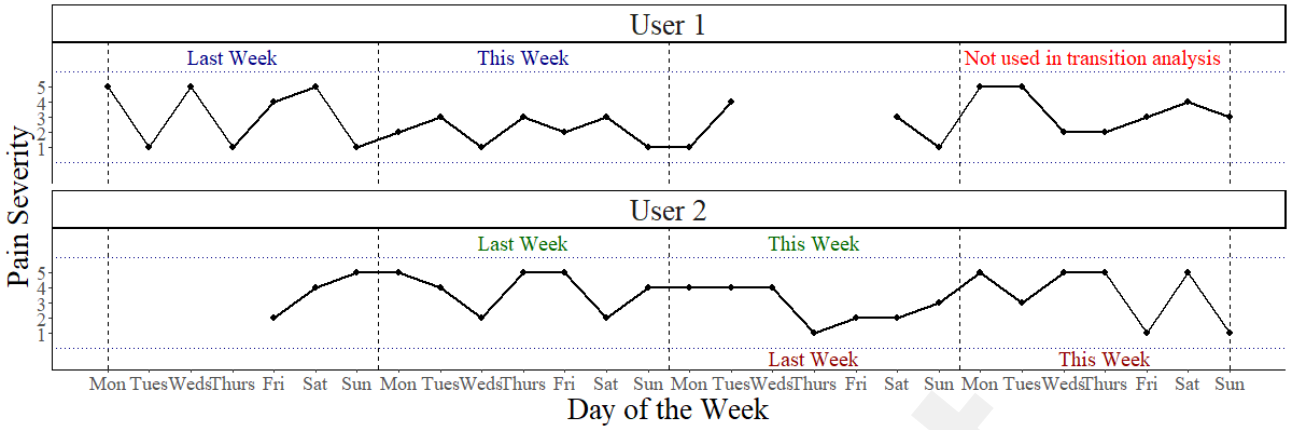


Figure 3: Example data from two participants, highlighting how their data is used to examine transitions between clusters. User 1 provided data in three complete participant-weeks. The first two are consecutive and therefore are used in the transition analysis. The final complete participant-week is not used. User 2 provided three complete participant-weeks. All three are used in the transition analysis. The middle week is labelled as both 'This Week' and 'Last Week' in different pairings.

Each trajectory was assigned a cluster in the CLARA program of the *k*-medoids cluster algorithm. The transition probabilities were then calculated as follows. For all trajectories in each cluster "Last week", the percentage of trajectories that transitioned to each cluster "This week" were calculated. The resulting percentages are reported in a transition matrix.

Data were analysed in R (version 4.1.2). Reporting of the analysis followed the STROBE guidelines [51].

Results

Data source

There were 2807 participants who contributed 21,919 participant-weeks of data to this analysis. The mean age was 51.2 (s.d. 12.8) years, and 2,333 participants (83%) were female. Table 1 reports the number of participants by age, sex, chronic pain condition and the average number of participant-weeks contributed to the analysis by members of the subgroup. Overall, older participants

contributed a greater number of participant-weeks than younger participants. Men contributed slightly more (8.1) participant-weeks than women (7.7). Participants with osteoarthritis (9.1) and unspecific arthritis (9.0) contributed the highest number of participant-weeks, and participants with chronic headache (6.0) contributed the fewest participant-weeks. Comorbid conditions are described in Supplementary Materials 1, Table S1 and Table S2.

Table 1: Demographic information of the participants who contributed to the analysis and the average number of participant-weeks contributed by each subgroup.

		Number of participants (%)	Mean number of weekly trajectories contributed by participants
Age group (years)			
	17–24	67 (2.4)	5.2
	25–34	255 (9.1)	5.6
	35–44	508 (18.1)	6.8
	45–54	755 (26.9)	7.5
	55–64	788 (28.1)	8.6
	65–86	434 (15.5)	9.9
Sex			
	Female	2333 (83.1)	7.7
	Male	474 (16.9)	8.1
Chronic pain condition[a]			
	Rheumatoid arthritis	548 (19.5)	7.7
	Osteoarthritis	975 (34.7)	9.1
	Spondyloarthropathy	254 (9.0)	7.6
	Gout	96 (3.4)	7.8
	Unspecific arthritis	1028 (36.6)	9.0
	Fibromyalgia	718 (25.6)	7.1
	Chronic headache	274 (9.8)	6.0
	Neuropathic pain	427 (15.2)	7.5
	Other/no medical diagnosis	668 (23.8)	6.9

[a]Percentages exceed 100% because participants could report multiple chronic-pain conditions.

Identifying optimal number of clusters

The results of the CLARA algorithm are shown below. Figure 4 reports the remaining variability within clusters as the "within-cluster sum of squares" at each value of *k*. There is an elbow at *k* = 4, suggesting that most of the observed variability can be explained by a solution with four clusters with diminishing returns for including further clusters in the solution. Four clusters reduced the within-cluster sum of squares from 159,100 to 66,507 and so the clustering algorithm describes 58.2% percent of the variability in the data. Each cluster contained more than 5% of the pain trajectories. Therefore, four clusters provide an appropriate choice for these data.

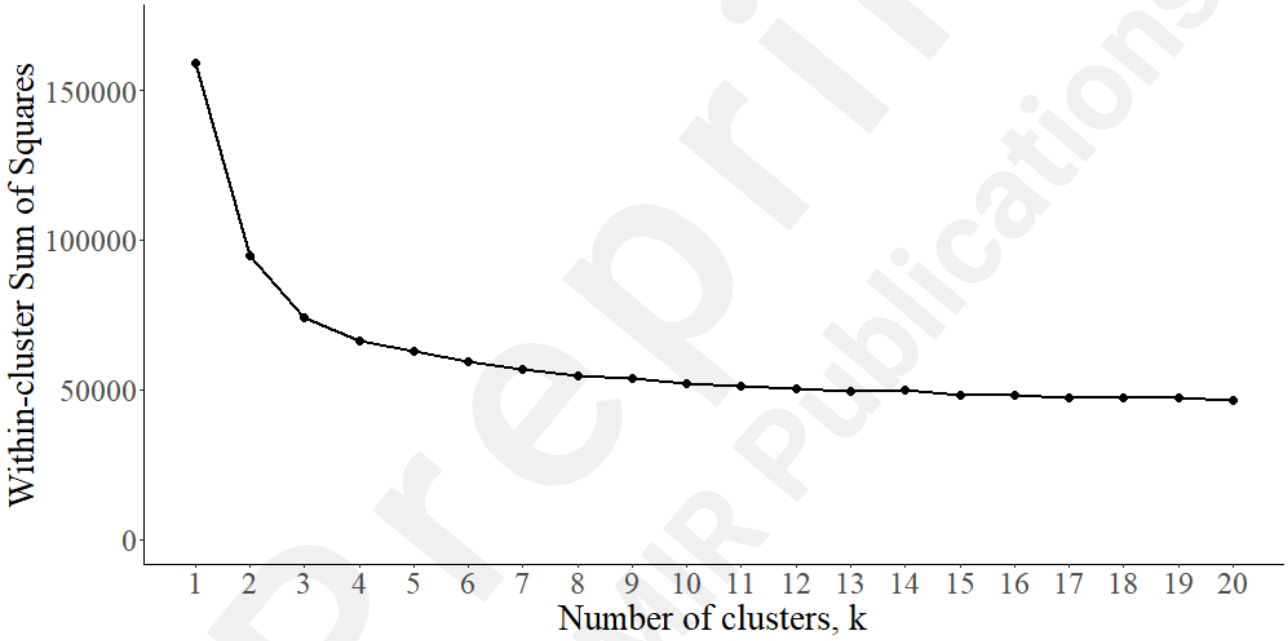


Figure 4: Unexplained variability across different cluster (*k*) solutions. The within-sum-of-squares indicates the remaining variance within clusters. An elbow at *k* = 4 suggests an appropriate solution, with diminishing returns for the inclusion of further clusters.

The trajectories in each cluster are shown by the spaghetti plot in Figure 5. Trajectories are weighted such that thicker lines represent a higher number of trajectories following the path. The red line represents the medoid of the *k*-medoids algorithm. The clusters can be named by examining the medoid. They represent A = "no or low pain", B = "mild pain", C = "moderate pain", D = "severe

pain". Cluster A contained 1714 (7.8%) trajectories, cluster B contained 8246 (37.6%) trajectories, cluster C contained 8376 (38.2%) trajectories, and cluster D contained 3583 (16.3%) trajectories.

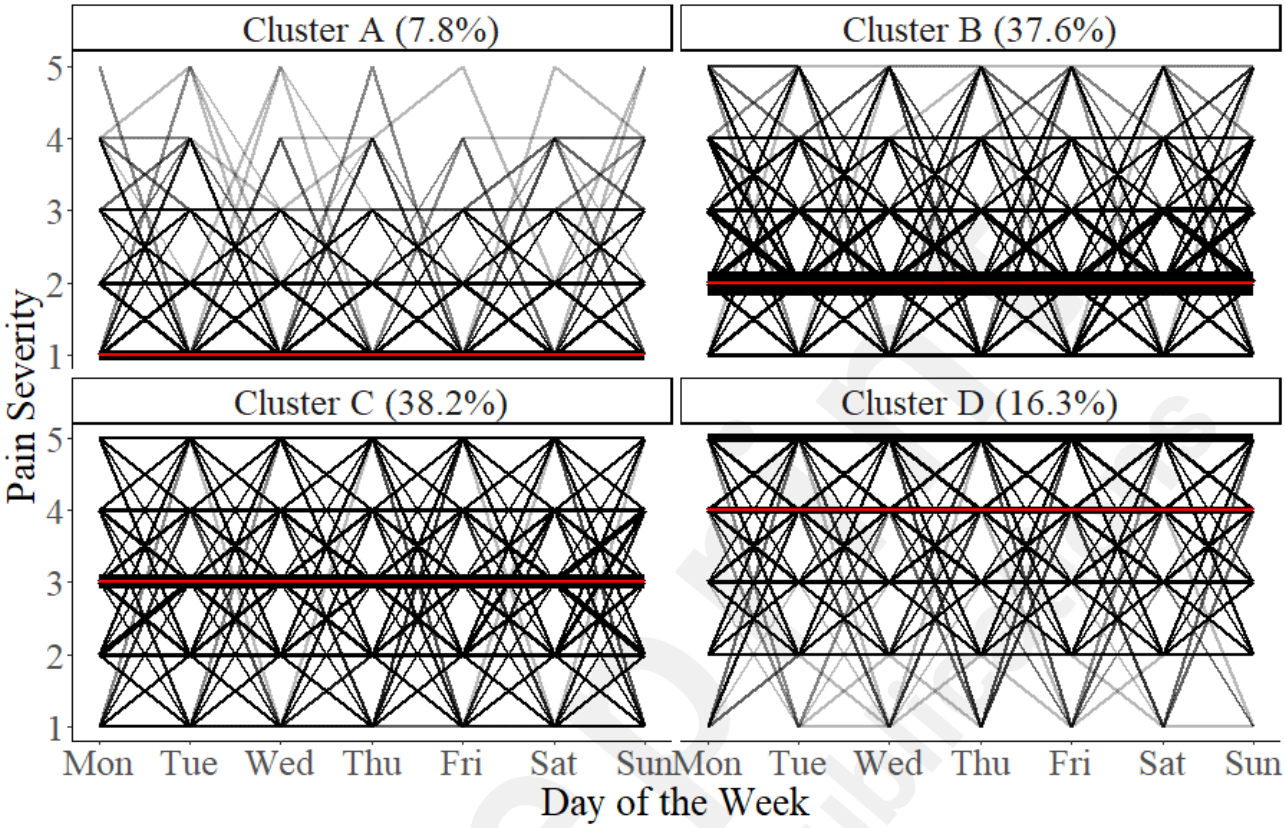


Figure 5: Weighted spaghetti plot of trajectories assigned to each cluster. The weight (and transparency) of each path represents the number of trajectories following that path. The red line represents the medoid of the cluster. Cluster A = "no or low pain", B = "mild pain", C = "moderate pain" and D = "severe pain". The percentage of trajectories assigned to each cluster is shown.

Sensitivity analyses

Sensitivity analysis 1 (KmL algorithm & Euclidean distance): Full results are presented in Supplementary Materials 2. The plot visualising within-cluster sum of squares against *k* for this analysis is similar to that of the main analysis and has an elbow at *k* = 4 (Figure 6). The optimal four-cluster solution describes 60.0% of the observed variability. The descriptions of the spaghetti plots (i.e. Cluster A = no or low pain, Cluster B = mild pain etc) are the same as the main analysis, despite use of a mean rather than medoid to describe the average trajectory in each cluster. In total, 18,895 (86.2%) trajectories were assigned to the same cluster as the main analysis, indicating similar results. Clusters B and C remain the largest clusters (8484 (38.7%) and 8001 (36.5%) trajectories, respectively), although Cluster A is larger in this sensitivity analysis than the main analysis (2493 (11.4%) in contrast to 1714 (7.8%)).

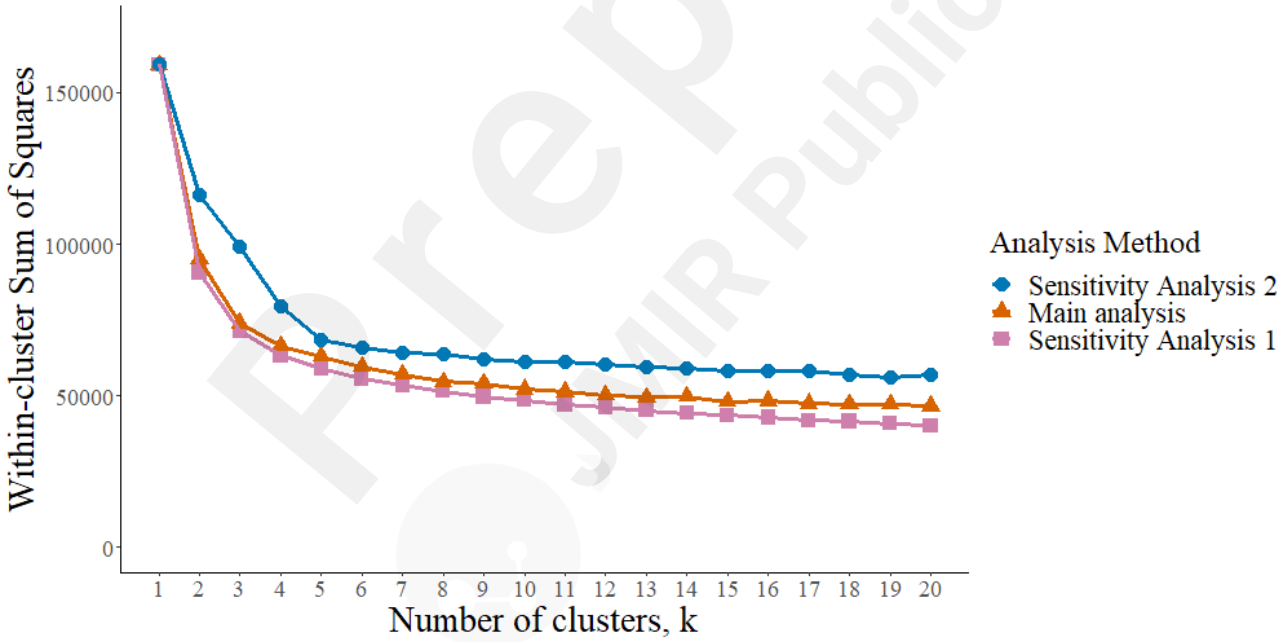


Figure 6: Unexplained variability across different cluster (*k*) solutions for the main analysis and two sensitivity analyses. In the main analysis and Sensitivity Analysis 1, there is an elbow at *k* = 4. In Sensitivity Analysis 2, there is an elbow at *k* = 5. Separate graphs are provided in Supplementary Materials 2.

Sensitivity analysis 2 (CLARA algorithm & Jaccard distance): Full results are presented in Supplementary Materials 2. The plot of *k* against WSS for this analysis has an elbow at *k* = 5 (Figure 6). However, one cluster contained 990 (4.5%) trajectories in the five-cluster solution, which did not meet the criteria for cluster sizes >5% and therefore a four-cluster solution remained optimal in this analysis. A four-cluster solution describes 50.0% of the variability. Spaghetti plots of the four-cluster solution show the same descriptions as the main analysis. In total, 20197 (92.1%) trajectories were assigned to the same cluster as in the main analysis.

Sensitivity analysis 3 (day of the week): Full results are presented in Supplementary Materials 3. Each plot of within-cluster sum of squares against *k* suggested an optimal solution at *k* = 4. The proportions of trajectories assigned to each cluster in each four-cluster solution are similar to the main analysis. The proportions in cluster A ranged between 7.7% and 7.9%, Cluster B between 37.5% and 37.7%, Cluster C between 38.0% and 38.4% and Cluster D between 16.2% and 16.4%. These results show that the main analysis is robust to the day of the week that the trajectories begin.

Description of clusters

The average proportion of time spent in different clusters across different characteristics (age, sex and condition) are summarised in Table 2. The oldest age bracket (65–86 years old) spent less time (9.8%) in the severe-pain cluster compared to the youngest age bracket (17–24, 28.3%). Women spent more time in the severe-pain cluster (18.0%) than men (12.3%) and less time in the lowest-pain cluster (Female: 6.5%, Male: 7.9%). Participants with fibromyalgia and neuropathic pain spent the most time in the severe-pain cluster (31.5% and 31.1%, respectively). Participants with rheumatoid arthritis spent the most time in the lowest-pain cluster (7.8%).

Table 2: Percentage of time spent in each cluster by baseline characteristics.

	Average percentage of time spent in cluster A (%)	Average percentage of time spent in cluster B (%)	Average percentage of time spent in cluster C (%)	Average percentage of time spent in cluster D (%)

All					
		6.7	36.8	39.4	17.0
Age group (years)					
	17–24	3.4	28.5	39.8	28.3
	25–34	6.7	32.0	41.5	19.8
	35–44	5.3	31.5	38.6	24.5
	45–54	5.6	36.3	39.4	18.7
	55–64	8.4	38.5	40.4	12.7
	65–86	7.9	44.9	37.3	9.8
Sex					
	Female	6.5	35.6	39.9	18.0
	Male	7.9	42.5	37.2	12.3
Chronic pain condition[a]					
	Rheumatoid arthritis	7.8	38.9	39.5	13.8
	Osteoarthritis	5.4	34.7	42.6	17.2
	Spondyloarthropathy	4.1	31.8	43.5	20.6
	Gout	6.1	33.0	41.6	19.3
	Unspecific arthritis	6.3	39.0	38.5	16.2
	Fibromyalgia	1.7	19.1	47.7	31.5
	Chronic headache	5.1	28.7	40.3	25.9
	Neuropathic pain	3.3	23.6	42.0	31.1
	Other/no medical diagnosis	6.7	36.8	39.4	17.0

For each characteristic, the average percentage of time spent in each cluster by members of the characteristic are reported. [a]Percentages exceed 100% because participants could report multiple chronic-pain conditions.

Transition between clusters

There were 12,267 pairs of participant weeks from 1761 participants used in the transition analysis. The demographic data are compared to those in the main analysis in Supplementary Materials 4. In general, a slightly higher proportion of older people contributed to the transition analysis than did to the cluster analysis. For example, there were 434 people (15.5% of participants) aged 65–86 in the main analysis but 300 (17.0% of participants) in the transition analysis. There are no other differences in the demographics of participants contributing to the main analysis and the transition analysis.

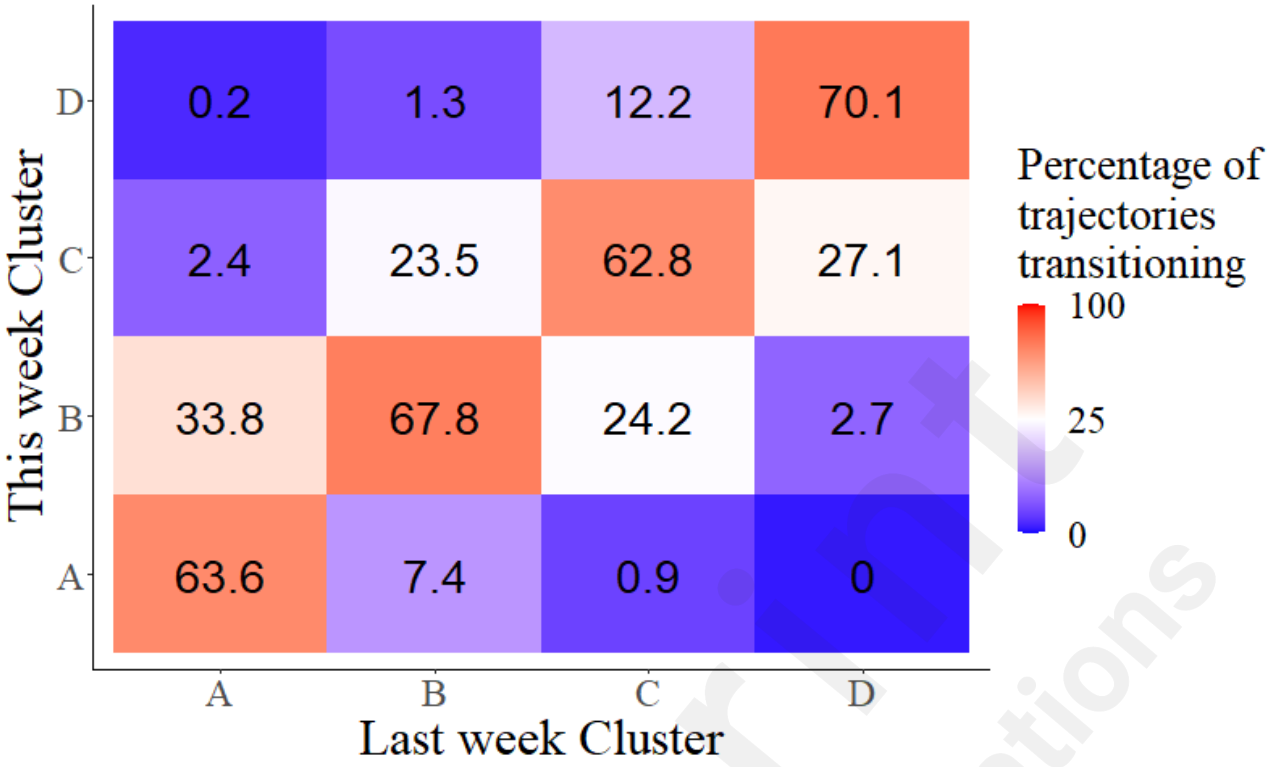


Figure 7: Transition matrix of movement between clusters on consecutive weeks. For membership in each cluster “Last Week”, the percentage of membership in each cluster “This Week” is shown. Random movement between clusters would suggest each combination has a transition percentage of 25%. Blue squares represent transitions that have a higher than random percentage (> 25%). Red squares represent transitions that have a lower than random percentage (< 25%). A white square would have exactly the random percentage (25%).

The percentages of consecutive trajectories transitioning between clusters are shown in Figure 7. For each cluster, the highest percentage of trajectories in the consecutive week remain in the same cluster, with the percentage values ranging between 63% and 70%. On average, 66% of trajectories remain in the same cluster. When individuals move between clusters, it is most frequently to an adjacent cluster. There are a very small percentage of consecutive weeks displaying movement between clusters two or more levels away.

Discussion

Principal Findings

This study identified and described clusters of weekly trajectories of pain severity in a large population-based mHealth study, to address four objectives in turn. First, we reported that four clusters (A = “no/low pain severity”, B = “mild pain”, C = “moderate pain” and D = “severe pain”) represented an optimal clustering solution for these data. In this solution, clusters B and C contained the greatest number of weekly pain trajectories.

Second, we conducted sensitivity analyses, to identify whether conclusions made about the first objective were robust to modified assumptions around the structure of the data. Two sensitivity analyses were conducted when the outcome variable was assumed to be (1) continuous and longitudinal and (2) categorical and not longitudinal. These analyses found that four clusters remained a suitable conclusion. A third sensitivity analysis found no differences in the clusters of trajectories starting on different days of the week.

Third, younger people and women contributed a greater number of trajectories to the severe-pain cluster when compared to older people and men, respectively. Participants with fibromyalgia and neuropathic pain contributed more trajectories to the severe-pain cluster than participants with other pain conditions, whereas participants with rheumatoid arthritis contributed more trajectories of to the no/low pain cluster than participants with other pain conditions.

Fourth, we examined transitions between clusters and found that about 66% of consecutive trajectories contributed to the same cluster. However, there was clear evidence of between-cluster movement with 34% of consecutive trajectories assigned to different clusters. Between-cluster movement was most likely to a neighbouring cluster; for example, moving from cluster 1 to cluster 2 was more common than moving from cluster 1 to cluster 3. This analysis demonstrates that overall, individuals tend to experience similar patterns of pain severity from week to week, although there are

substantial experiences of increase or decreases in pain severity, thereby reflecting the lived experience of people with chronic pain having variability in symptoms, and noting how pain can fluctuate between weeks.

People with chronic pain have highlighted a need to describe and predict the variability in the severity of their pain. Through clustering, this study has described four common experiences of pain severity, accounting for two-thirds of the observed variability. However, trajectories within each cluster are not homogenous and there remains within-cluster variation. To better describe the individual weekly pain experience, future work should explore the remaining variability within clusters.

Comparison with Prior Work

Many studies have identified clusters of pain trajectories among individuals living with chronic pain. Some have focussed on participants with one chronic pain condition, such as osteoarthritis [15, 36–39, 52–58], low back pain [13, 14, 27, 59–65], other back pain [25, 32, 49, 66], neck or shoulder pain [33, 61, 67, 68], leg pain [29], knee pain [69] or foot pain [70], whereas others have identified clusters among a broader population such as those with musculoskeletal pain [26, 31, 47, 71, 72] or general pain [48, 73–75]. Clusters in these studies were described by the severity of pain (e.g., no pain, very low pain, mild pain, moderate pain, high pain, severe pain), the level of change in pain severity (e.g., persistent, ongoing, episodic, worsening, recovering, fluctuating) or a combination of these features.

These previous studies have often considered only sparse data, with relatively large time intervals between consecutive data points. Of those gathering data for at least one year (27 studies), data were collected more than twelve times in only two studies [13, 67]. In these two studies, data were collected weekly for one year to explore the course of specific pain conditions (neck pain, low back pain). Kongsted et al. [13] used 12 models to identify between 5 and 12 clusters in each model.

Clusters were described by the severity of pain (e.g., “moderate”, “severe”) and also by temporal features of the trajectories (e.g., “episodic”, “recovery”, “ongoing”). Pico-Espinosa et al. [67] identified six clusters of pain described as “small improvement”, “moderate improvement”, “persistent”, “large improvement”, “slightly fluctuating”, “highly fluctuating”. The clusters identified in our analysis were described by the severity of pain, similar to clusters in studies with sparse data. Our clusters were unlikely to identify long-term disease development, as with trajectories over longer periods.

Similar to our study, some previous studies have used methods from the *k*-means/*k*-medoids family of algorithms. Knecht et al. [25] used the *KmL* package to identify two clusters in a ‘responders’ and ‘non-responders’ group. Weng et al. [76] used a *k*-median algorithm to identify four clusters of pain severity described by ‘slightly rise’, ‘completely drop’, ‘sudden rise’, and ‘steady group’. These studies both identified trajectories with changing pain while our study identified trajectories of weekly pain where the medoid was stable across the week.

In all studies listed above, the experiences of individuals were described by a single trajectory across the full duration of follow-up, whereas our study examined week-to-week transition between clusters. Kongsted et al. [77] have previously examined week-to-week pain severity across a year, using pre-defined clusters. They identified that 41% and 21% of respondents in two different datasets had stable pain over a year, defined as pain within 1 point of the mean pain value on an 11-point numerical rating scale. The remaining pain trajectories were classified as having a single episode of pain, being episodic or fluctuating. The transitions identified in our study suggest stability between 66% of consecutive weeks. However, some individuals in our study may experience the other longer-term descriptions from Kongsted et al. For example, an individual might not transition out of a cluster for most of the year yet experience a single episode. Future studies should further examine the movement between different pain states and identify the drivers of these transitions.

Strengths and Limitations

A number of strengths and limitations of this study should be considered. First, a strength was that participants could contribute daily data for up to 64 weeks. This frequent and granular data collection, enabled by mHealth, overcame limitations of sparse data collection in previous studies (as identified by Beukenhorst et al. [78]). As a result, this study was able to analyse the weekly trajectories contributed by participants, determining common pain patterns among a chronic pain population at a more granular scale than previously investigated.

Second, the analysis presented in this paper modelled weekly pain trajectories rather than individual people. In contrast to studies that assign each individual to a single cluster across the whole follow-up, individuals were able to transition between different pain clusters over time as their pain experience changed and their condition developed. These transitions were observed in 34% of consecutive weeks, and this flexibility can be used in future work to explore the mechanisms driving movement between clusters.

Third, assumptions about the ordinal, longitudinal form of the data were modified in sensitivity analyses. A four-cluster solution was most suitable for each analysis, indicating a strong model-independent signal in the data and a more robust conclusion about the most suitable number of clusters. Furthermore, the assignment of trajectories to each cluster were similar in each analysis (at least 86% similarity), indicating further stability in the results. There were benefits to the use of both the *k*-medoids algorithm and longitudinal adaptation of the *k*-means algorithm used in this analysis. First, neither of these methods require parametric assumptions about the form of the data [50]. Second, no prior assumptions, including the shape of the trajectory, are required by the algorithms [79]. Therefore, this data-driven approach made limited assumptions about the form of the data.

There were also limitations to the study. The data used in this study were from a population-based study which represented the UK population. Cloudy with a Chance of Pain recruited participants from all UK postcodes, although men and those in the age brackets 17–34 and 75+ were

underrepresented in the study population [24]. Despite being a smaller population, older people and men contributed more trajectories on average and were more likely to contribute trajectories to a less severe pain category. As these clusters will be used in the development of a pain-forecasting model, clusters should be generalisable to the chronic-pain population, and there remains the possibility that different pain clusters and between-cluster transitions could be realized among those who did and did not contribute to the study. Although it is unlikely that our large study population would display different pain clusters and transitions to the chronic-pain population that would use a smartphone tracking application, it remains a possibility that should be explored in future studies.

This analysis further selected participants by the requirement to provide a week of complete pain-severity data, thus excluding missing data. There are reasons that data might be missing not at random, including missing due to severe pain, missing due to low pain severity, and missing due to stable pain that results in repetitive score input and thus disengagement. The transition analysis also further selected participants by requiring two weeks of complete pain-severity data. However, the age, sex and chronic-pain conditions of respondents in the main analysis and transition analysis (Table 4) were similar to those in the full-study population (Supplementary Table 1 in Dixon et al. [16]), suggesting that the included participants were representative of the study population.

There were limitations in the method used for clustering. First, the absence of parametric assumptions in either the *k*-medoids algorithm or KmL package resulted in goodness-of-fit measures being inappropriate [79]. Therefore, the elbow method was used to select the optimum number of clusters. However, the use of the elbow method introduces subjectivity. Second, both the *k*-medoids algorithm and the KmL package require random starting values for the cluster centres, which can add volatility to the results. This volatility was mitigated by repeating the algorithms 20 times each and selecting the solution with the lowest remaining variability within clusters.

Conclusions

Previous research has highlighted a need to better understand pain variability experienced by individuals with chronic pain [20]. Feelings of uncertainty among people with chronic pain have led them to want to better understand the pain that they may experience in the future. Clustering weekly pain trajectories offers a first step to better understanding common experiences of pain severity. Once these common experiences are better described, they can be used in future work to predict movement between clusters.

There are limited methods available for clustering pain severity that respect the ordinal and longitudinal nature of patient-generated health data in a computationally inexpensive manner. The clustering method and subsequent sensitivity analyses presented in this paper suggest that the use of *k*-medoids is robust to assumptions about the data structure.

This study has identified four distinct patterns of weekly pain severity: no or low pain, mild pain, moderate pain or severe pain. These can be used to describe short-term pain experiences of people with chronic pain. Future work is required to identify how these clusters can be used in a forecasting model of pain. First, there remains individual variability within clusters of pain severity. Participants of PPI have identified that fluctuations in pain severity should be forecasted, and therefore within-cluster variability should be quantified to further understand the weekly pain experience of individuals. Second, transition of individuals between clusters should be explored to identify drivers of movement between pain clusters on an individual level. The clusters identified in this study and future work to understand within-cluster variability and drivers of movement between clusters will enable a future pain-forecasting model.

Acknowledgements:

Conflicts of interest:

WGD has received consultancy fees from Google, and DMS has received consultancy fees from Palta, both unrelated to this work. All other authors have no conflicts of interest to declare.

Funding:

This study was supported by the Centre for Epidemiology Versus Arthritis (grant number 21755). Cloudy with a Chance of Pain was funded by Versus Arthritis (grant reference 21225), with additional support from the Centre for Epidemiology (grants 21755 and 20380). TH receives funding from the Royal Society (grant number INF/R2/180067) and the Alan Turing Institute for Data Science and Artificial Intelligence. The funders of the study had no role in study design, data interpretation, writing of the manuscript, or the decision to submit for publication.

Data availability statement:

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Code Availability:

Data management and analyses were performed in R 4.1.2. Code is available on reasonable request.

Reference List

1. McQueenie R, Jani BD, Siebert S, McLoone P, McCowan C, Macdonald S, Mair FS, Nicholl BI (2021) Prevalence of chronic pain in LTCs and multimorbidity: A cross-sectional study using UK Biobank. J Multimorb Comorbidity 11:263355652110058. https://doi.org/10.1177/26335565211005870

2. Wolfe F, Clauw DJ, Fitzcharles M-A, Goldenberg DL, Katz RS, Mease P, Russell AS, Russell IJ, Winfield JB, Yunus MB (2010) The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res 62:600–610. https://doi.org/10.1002/acr.20140

3. Steineck A, Bradford MC, O'Daffer A, Fladeboe KM, O'Donnell MB, Scott S, Yi-Frazier JP, Rosenberg AR (2022) Quality of Life in Adolescents and Young Adults: The Role of Symptom Burden. J Pain Symptom Manage S0885-3924(22)00742–4. https://doi.org/10.1016/j.jpainsymman.2022.05.017

4. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J (2005) Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 113:9–19. https://doi.org/10.1016/j.pain.2004.09.012

5. Schneider S, Junghaenel DU, Keefe FJ, Schwartz JE, Stone AA, Broderick JE (2012) Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. Pain 153:813–822. https://doi.org/10.1016/j.pain.2012.01.001

6. Wesolowicz DM, Bishop MD, Robinson ME (2021) An Examination of Day-to-Day and Intraindividual Pain Variability in Low Back Pain. Pain Med Malden Mass 22:2263–2275. https://doi.org/10.1093/pm/pnab119

7. Crowe M, Whitehead L, Seaton P, Jordan J, Mccall C, Maskill V, Trip H (2017) Qualitative meta-synthesis: the experience of chronic pain across conditions. J Adv Nurs 73:1004–1016. https://doi.org/10.1111/jan.13174

8. Johnson LM, Zautra AJ, Davis MC (2006) The role of illness uncertainty on coping with fibromyalgia symptoms. Health Psychol 25:696–703. https://doi.org/10.1037/0278-6133.25.6.696

9. Edwards RR, Dworkin RH, Turk DC, Angst MS, Dionne R, Freeman R, Hansson P, Haroutounian S, Arendt-Nielsen L, Attal N, Baron R, Brell J, Bujanover S, Burke LB, Carr D, Chappell AS, Cowan P, Etropolski M, Fillingim RB, Gewandter JS, Katz NP, Kopecky EA, Markman JD, Nomikos G, Porter L, Rappaport BA, Rice ASC, Scavone JM, Scholz J, Simon LS, Smith SM, Tobias J, Tockarshewsky T, Veasley C, Versavel M, Wasan AD, Wen W, Yarnitsky D (2016) Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. Pain 157:1851–1871. https://doi.org/10.1097/j.pain.0000000000000602

10. Tupper SM, Rosenberg AM, Pahwa P, Stinson JN (2013) Pain Intensity Variability and Its Relationship With Quality of Life in Youths With Juvenile Idiopathic Arthritis. Arthritis Care

Res 65:563–570. https://doi.org/10.1002/acr.21850

11. Zakoscielna KM, Parmelee PA (2013) Pain variability and its predictors in older adults: depression, cognition, functional status, health, and pain. J Aging Health 25:1329–1339. https://doi.org/10.1177/0898264313504457

12. Stone AA, Broderick JE, Goldman RE, Junghaenel DU, Bolton A, May M, Schneider S (2021) I. Indices of Pain Intensity Derived From Ecological Momentary Assessments: Rationale and Stakeholder Preferences. J Pain 22:359–370. https://doi.org/10.1016/j.jpain.2020.08.003

13. Kongsted A, Kent P, Hestbaek L, Vach W (2015) Patients with low back pain had distinct clinical course patterns that were typically neither complete recovery nor constant pain. A latent class analysis of longitudinal data. Spine J Off J North Am Spine Soc 15:885–894. https://doi.org/10.1016/j.spinee.2015.02.012

14. Dunn KM, Jordan K, Croft PR (2006) Characterizing the Course of Low Back Pain: A Latent Class Analysis. Am J Epidemiol 163:754–761. https://doi.org/10.1093/aje/kwj100

15. Bastick AN, Wesseling J, Damen J, Verkleij SPJ, Emans PJ, Bindels PJE, Bierma-Zeinstra SMA (2016) Defining knee pain trajectories in early symptomatic knee osteoarthritis in primary care: 5-year results from a nationwide prospective cohort study (CHECK). Br J Gen Pract J R Coll Gen Pract 66:e32-39. https://doi.org/10.3399/bjgp15X688129

16. Dixon WG, Beukenhorst AL, Yimer BB, Cook L, Gasparrini A, El-Hay T, Hellman B, James B, Vicedo-Cabrera AM, Maclure M, Silva R, Ainsworth J, Pisaniello HL, House T, Lunt M, Gamble C, Sanders C, Schultz DM, Sergeant JC, McBeth J (2019) How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digit Med 2:105. https://doi.org/10.1038/s41746-019-0180-3

17. Check Hayden E (2016) Mobile-phone health apps deliver data bounty. Nature 531:422–423. https://doi.org/10.1038/531422a

18. Shiffman S, Stone AA, Hufford MR (2008) Ecological Momentary Assessment. Annu Rev Clin Psychol 4:1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

19. Rahman QA, Janmohamed T, Pirbaglou M, Clarke H, Ritvo P, Heffernan JM, Katz J (2018) Defining and Predicting Pain Volatility in Users of the Manage My Pain App: Analysis Using Data Mining and Machine Learning Methods. J Med Internet Res 20:e12001. https://doi.org/10.2196/12001

20. Little CL, Druce KL, Dixon WG, Schultz DM, House T, McBeth J (2023) What do people living with chronic pain want from a pain forecast? A research prioritisation study. Epidemiology

21. Flurey CA, Morris M, Richards P, Hughes R, Hewlett S (2014) It's like a juggling act: rheumatoid arthritis patient perspectives on daily life and flare while on current treatment regimes. Rheumatology 53:696–703. https://doi.org/10.1093/rheumatology/ket416

22. Liddell TM, Kruschke JK (2018) Analyzing ordinal data with metric models: What could possibly go wrong? J Exp Soc Psychol 79:328–348. https://doi.org/10.1016/j.jesp.2018.08.009

23. Schultz DM, Beukenhorst AL, Yimer BB, Cook L, Pisaniello HL, House T, Gamble C, Sergeant JC, McBeth J, Dixon WG (2020) Weather Patterns Associated with Pain in Chronic-Pain Sufferers. Bull Am Meteorol Soc 101:E555–E566. https://doi.org/10.1175/BAMS-D-19-0265.1

24. Druce KL, McBeth J, van der Veer SN, Selby DA, Vidgen B, Georgatzis K, Hellman B, Lakshminarayana R, Chowdhury A, Schultz DM, Sanders C, Sergeant JC, Dixon WG (2017) Recruitment and Ongoing Engagement in a UK Smartphone Study Examining the Association Between Weather and Pain: Cohort Study. JMIR MHealth UHealth 5:e168. https://doi.org/10.2196/mhealth.8162

25. Knecht C, Hartnack S, Sick B, Riner F, Schweinhardt P, Wirth B (2020) A prospective observational study on trajectories and prognostic factors of mid back pain. BMC Musculoskelet Disord 21:554. https://doi.org/10.1186/s12891-020-03534-5

26. Crook J, Moldofsky H (1996) The clinical course of musculoskeletal pain in empirically derived groupings of injured workers. Pain 67:427–433. https://doi.org/10.1016/0304-3959(96)03151-x

27. Axén I, Bodin L, Bergström G, Halasz L, Lange F, Lövgren PW, Rosenbaum A, Leboeuf-Yde C, Jensen I (2011) Clustering patients on the basis of their individual course of low back pain over a six month period. BMC Musculoskelet Disord 12:99. https://doi.org/10.1186/1471-2474-12-99

28. Brennan PL, SooHoo S (2020) Effects of Mental Health Disorders on Nursing Home Residents' Nine-Month Pain Trajectories. Pain Med Malden Mass 21:488–500. https://doi.org/10.1093/pm/pnz177

29. Ogollah RO, Konstantinou K, Stynes S, Dunn KM (2018) Determining One-Year Trajectories of Low-Back-Related Leg Pain in Primary Care Patients: Growth Mixture Modeling of a Prospective Cohort Study. Arthritis Care Res 70:1840–1848. https://doi.org/10.1002/acr.23556

30. Battaglia M, Garon-Carrier G, Brendgen M, Feng B, Dionne G, Vitaro F, Tremblay RE, Boivin M (2020) Trajectories of pain and anxiety in a longitudinal cohort of adolescent twins. Depress Anxiety 37:475–484. https://doi.org/10.1002/da.22992

31. Neupane S, Nygård C-H, Prakash KC, von Bonsdorff MB, von Bonsdorff ME, Seitsamo J, Rantanen T, Ilmarinen J, Leino-Arjas P (2018) Multisite musculoskeletal pain trajectories from midlife to old age: a 28-year follow-up of municipal employees. Occup Environ Med 75:863–870. https://doi.org/10.1136/oemed-2018-105235

32. Dunn KM, Campbell P, Jordan KP (2013) Long-term trajectories of back pain: cohort study with 7-year follow-up. BMJ Open 3:e003838. https://doi.org/10.1136/bmjopen-2013-003838

33. Hallman DM, Rasmussen CDN, Jørgensen MB, Holtermann A (2018) Time course of neck-shoulder pain among workers: A longitudinal latent class growth analysis. Scand J Work Environ Health 44:47–57. https://doi.org/10.5271/sjweh.3690

34. Vermunt JK (2008) Latent class and finite mixture models for multilevel data sets. Stat Methods Med Res 17:33–51. https://doi.org/10.1177/0962280207081238

35. Hebert JJ, Abraham E, Wedderkopp N, Bigney E, Richardson E, Darling M, Hall H, Fisher CG,

Rampersaud YR, Thomas KC, Jacobs B, Johnson M, Paquet J, Attabib N, Jarzem P, Wai EK, Rasoulinejad P, Ahn H, Nataraj A, Stratton A, Manson N (2019) Patients undergoing surgery for lumbar spinal stenosis experience unique courses of pain and disability: A group-based trajectory analysis. PloS One 14:e0224200. https://doi.org/10.1371/journal.pone.0224200

36. Wieczorek M, Rotonda C, Coste J, Pouchot J, Saraux A, Guillemin F, Rat A-C (2020) Trajectory analysis combining pain and physical function in individuals with knee and hip osteoarthritis: results from the French KHOALA cohort. Rheumatol Oxf Engl 59:3488–3498. https://doi.org/10.1093/rheumatology/keaa148

37. Dai Z, Lu N, Niu J, Felson DT, Zhang Y (2017) Dietary Fiber Intake in Relation to Knee Pain Trajectory. Arthritis Care Res 69:1331–1339. https://doi.org/10.1002/acr.23158

38. Radojčić MR, Arden NK, Yang X, Strauss VY, Birrell F, Cooper C, Kluzek S, the VIDEO Trial Investigators (2020) Pain trajectory defines knee osteoarthritis subgroups: a prospective observational study. Pain 161:2841–2851. https://doi.org/10.1097/j.pain.0000000000001975

39. Pan F, Tian J, Aitken D, Cicuttini F, Jones G (2018) Predictors of pain severity trajectory in older adults: a 10.7-year follow-up study. Osteoarthritis Cartilage 26:1619–1626. https://doi.org/10.1016/j.joca.2018.08.002

40. Pan F, Tian J, Cicuttini F, Jones G (2020) Metabolic syndrome and trajectory of knee pain in older adults. Osteoarthritis Cartilage 28:45–52. https://doi.org/10.1016/j.joca.2019.05.030

41. Kassambara A (2017) Practical guide to cluster analysis in R: unsupervised machine learning, Edition 1. STHDA, Erscheinungsort nicht ermittelbar

42. Kaufman L, Rousseeuw PJ (1986) CLUSTERING LARGE DATA SETS. In: Pattern Recognition in Practice. Elsevier, pp 425–437

43. Kaufman L, Rousseeuw PJ (2005) Finding groups in data: an introduction to cluster analysis. Wiley, Hoboken, N.J

44. Sarang P (2023) Centroid-Based Clustering: Clustering Algorithms for Hard Clustering. In: Thinking Data Science. Springer International Publishing, Cham, pp 171–183

45. Hardy A (1996) On the number of clusters. Comput Stat Data Anal 23:83–96. https://doi.org/10.1016/S0167-9473(96)00022-9

46. Thomas MJ, Yu D, Nicholls E, Bierma-Zeinstra S, Conaghan PG, Stoner KJ, Neogi T, Parry EL, Peat G (2020) Short-Term Recovery Trajectories of Acute Flares in Knee Pain: A UK-Netherlands Multicenter Prospective Cohort Analysis. Arthritis Care Res 72:1687–1692. https://doi.org/10.1002/acr.24088

47. Airila A, Hakanen JJ, Luukkonen R, Lusa S, Punakallio A, Leino-Arjas P (2014) Developmental trajectories of multisite musculoskeletal pain and depressive symptoms: the effects of job demands and resources and individual factors. Psychol Health 29:1421–1441. https://doi.org/10.1080/08870446.2014.945929

48. James RJE, Walsh DA, Ferguson E (2018) General and disease-specific pain trajectories as predictors of social and political outcomes in arthritis and cancer. BMC Med 16:51.

https://doi.org/10.1186/s12916-018-1031-9

49. Canizares M, Rampersaud YR, Badley EM (2019) Course of Back Pain in the Canadian Population: Trajectories, Predictors, and Outcomes. Arthritis Care Res 71:1660–1670. https://doi.org/10.1002/acr.23811

50. Genolini C, Falissard B (2011) Kml: A package to cluster longitudinal data. Comput Methods Programs Biomed 104:e112–e121. https://doi.org/10.1016/j.cmpb.2011.05.008

51. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull World Health Organ 85:867–872. https://doi.org/10.2471/blt.07.045120

52. Verkleij SPJ, Hoekstra T, Rozendaal RM, Waarsing JH, Koes BW, Luijsterburg PAJ, Bierma-Zeinstra SMA (2012) Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. Ann Rheum Dis 71:1517–1523. https://doi.org/10.1136/annrheumdis-2011-200687

53. Wesseling J, Bastick AN, ten Wolde S, Kloppenburg M, Lafeber FPJG, Bierma-Zeinstra SMA, Bijlsma JWJ (2015) Identifying Trajectories of Pain Severity in Early Symptomatic Knee Osteoarthritis: A 5-year Followup of the Cohort Hip and Cohort Knee (CHECK) Study. J Rheumatol 42:1470–1477. https://doi.org/10.3899/jrheum.141036

54. Nicholls E, Thomas E, van der Windt DA, Croft PR, Peat G (2014) Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthritis Cartilage 22:2041–2050. https://doi.org/10.1016/j.joca.2014.09.026

55. Bastick AN, Verkleij SPJ, Damen J, Wesseling J, Hilberdink WKHA, Bindels PJE, Bierma-Zeinstra SMA (2016) Defining hip pain trajectories in early symptomatic hip osteoarthritis – 5 year results from a nationwide prospective cohort study (CHECK). Osteoarthritis Cartilage 24:768–775. https://doi.org/10.1016/j.joca.2015.11.023

56. Collins JE, Katz JN, Dervan EE, Losina E (2014) Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 22:622–630. https://doi.org/10.1016/j.joca.2014.03.009

57. Johnson AJ, Vasilopoulos T, Booker SQ, Cardoso J, Terry EL, Powell-Roach K, Staud R, Kusko DA, Addison AS, Redden DT, Goodin BR, Fillingim RB, Sibille KT (2021) Knee pain trajectories over 18 months in non-Hispanic Black and non-Hispanic White adults with or at risk for knee osteoarthritis. BMC Musculoskelet Disord 22:415. https://doi.org/10.1186/s12891-021-04284-8

58. Atukorala I, Downie A, Pathmeswaran A, Deveza LMA, Chang T, Zhang Y, Hunter DJ (2022) Short-term pain trajectories in patients with knee osteoarthritis. Int J Rheum Dis 25:281–294. https://doi.org/10.1111/1756-185X.14267

59. Tamcan O, Mannion AF, Eisenring C, Horisberger B, Elfering A, Müller U (2010) The course of chronic and recurrent low back pain in the general population. Pain 150:451–457. https://doi.org/10.1016/j.pain.2010.05.019

60. Coenen P, Smith A, Paananen M, O'Sullivan P, Beales D, Straker L (2017) Trajectories of Low Back Pain From Adolescence to Young Adulthood. Arthritis Care Res 69:403–412. https://doi.org/10.1002/acr.22949

61. Ailliet L, Rubinstein SM, Hoekstra T, van Tulder MW, de Vet HCW (2018) Long-term trajectories of patients with neck pain and low back pain presenting to chiropractic care: A latent class growth analysis. Eur J Pain Lond Engl 22:103–113. https://doi.org/10.1002/ejp.1094

62. Downie AS, Hancock MJ, Rzewuska M, Williams CM, Lin C-WC, Maher CG (2016) Trajectories of acute low back pain: a latent class growth analysis. Pain 157:225–234. https://doi.org/10.1097/j.pain.0000000000000351

63. Panken G, Hoekstra T, Verhagen A, van Tulder M, Twisk J, Heymans MW (2016) Predicting chronic low-back pain based on pain trajectories in patients in an occupational setting: an exploratory analysis. Scand J Work Environ Health 42:520–527. https://doi.org/10.5271/sjweh.3584

64. Macedo LG, Maher CG, Latimer J, McAuley JH, Hodges PW, Rogers WT (2014) Nature and determinants of the course of chronic low back pain over a 12-month period: a cluster analysis. Phys Ther 94:210–221. https://doi.org/10.2522/ptj.20120416

65. Kyrönlahti SM, Nygård C-H, K C P, Neupane S (2022) Trajectories of low back pain from midlife to retirement and functional ability at old age. Eur J Public Health 32:497–503. https://doi.org/10.1093/eurpub/ckab191

66. Enthoven WTM, Koes BW, Bierma-Zeinstra SMA, Bueving HJ, Bohnen AM, Peul WC, van Tulder MW, Berger MY, Luijsterburg PAJ (2016) Defining trajectories in older adults with back pain presenting in general practice. Age Ageing 45:878–883. https://doi.org/10.1093/ageing/afw127

67. Pico-Espinosa OJ, Côté P, Hogg-Johnson S, Jensen I, Axén I, Holm LW, Skillgate E (2019) Trajectories of Pain Intensity Over 1 Year in Adults With Disabling Subacute or Chronic Neck Pain: Clin J Pain 35:678–685. https://doi.org/10.1097/AJP.0000000000000727

68. Eriks-Hoogland IE, Hoekstra T, de Groot S, Stucki G, Post MW, van der Woude LH (2014) Trajectories of musculoskeletal shoulder pain after spinal cord injury: Identification and predictors. J Spinal Cord Med 37:288–298. https://doi.org/10.1179/2045772313Y.0000000168

69. Haybatollahi SM, James RJE, Fernandes G, Valdes A, Doherty M, Zhang W, Walsh DA, Ferguson E (2022) Identifying multiple knee pain trajectories and the prediction of opioid and NSAID medication used: A latent class growth approach. Pain Pract Off J World Inst Pain 22:210–221. https://doi.org/10.1111/papr.13082

70. Marshall M, Blagojevic-Bucknall M, Rathod-Mistry T, Thomas MJ, Edwards JJ, Peat G, Menz HB, Roddy E (2023) Identifying Long-Term Trajectories of Foot Pain Severity and Potential Prognostic Factors: A Population-Based Cohort Study. Arthritis Care Res 75:1123–1131. https://doi.org/10.1002/acr.24823

71. Leino-Arjas P, Rajaleid K, Mekuria G, Nummi T, Virtanen P, Hammarström A (2018) Trajectories of musculoskeletal pain from adolescence to middle age: the role of early depressive symptoms, a 27-year follow-up of the Northern Swedish Cohort. Pain 159:67–74.

https://doi.org/10.1097/j.pain.0000000000001065

72. Oakman J, Neupane S, Kyrönlahti S, Nygård C-H, Lambert K (2022) Musculoskeletal pain trajectories of employees working from home during the COVID-19 pandemic. Int Arch Occup Environ Health 95:1891–1901. https://doi.org/10.1007/s00420-022-01885-1

73. Neupane S, Lallukka T, Pietiläinen O, Rahkonen O, Leino-Arjas P (2020) Trajectories of multisite musculoskeletal pain in midlife: Associations with common mental disorders. Eur J Pain Lond Engl 24:364–373. https://doi.org/10.1002/ejp.1492

74. Sun F, Park NS, Wardian J, Lee BS, Roff LL, Klemmack DL, Parker MW, Koenig HG, Sawyer PL, Allman RM (2013) Predicting the Trajectories of Perceived Pain Intensity in Southern Community-Dwelling Older Adults: The Role of Religiousness. Res Aging 35:. https://doi.org/10.1177/0164027512456402

75. Wickrama KAS, Lee TK, O'Neal CW (2021) Heterogeneity in Husbands' and Wives' Physical Pain Trajectories Over Mid-Later Years: Biopsychosocial Stratification and Implications for Later-Life Well-Being. The Gerontologist 61:930–941. https://doi.org/10.1093/geront/gnab007

76. Weng Y, Tian L, Tedesco D, Desai K, Asch SM, Carroll I, Curtin C, McDonald KM, Hernandez-Boussard T (2020) Trajectory analysis for postoperative pain using electronic health records: A nonparametric method with robust linear regression and K-medians cluster analysis. Health Informatics J 26:1404–1418. https://doi.org/10.1177/1460458219881339

77. Kongsted A, Hestbæk L, Kent P (2017) How can latent trajectories of back pain be translated into defined subgroups? BMC Musculoskelet Disord 18:285. https://doi.org/10.1186/s12891-017-1644-8

78. Beukenhorst AL, Schultz DM, McBeth J, Sergeant JC, Dixon WG (2020) Are weather conditions associated with chronic musculoskeletal pain? Review of results and methodologies. Pain 161:668–683. https://doi.org/10.1097/j.pain.0000000000001776

79. Genolini C, Alacoque X, Sentenac M, Arnaud C (2015) **kml** and **kml3d** : *R* Packages to Cluster Longitudinal Data. J Stat Softw 65:. https://doi.org/10.18637/jss.v065.i04

Reference List

1. McQueenie R, Jani BD, Siebert S, McLoone P, McCowan C, Macdonald S, et al. Prevalence of chronic pain in LTCs and multimorbidity: A cross-sectional study using UK Biobank. J Multimorb Comorbidity 2021;11. doi:10.1177/26335565211005870

2. Wolfe F, Clauw DJ, Fitzcharles M-A, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res 2010;62(5):600–610. doi:10.1002/acr.20140

3. Steineck A, Bradford MC, O'Daffer A, Fladeboe KM, O'Donnell MB, Scott S, et al. Quality of Life in Adolescents and Young Adults: The Role of Symptom Burden. J Pain Symptom Manage 2022 Sep;64(3):244–253.e2. doi:10.1016/j.jpainsymman.2022.05.017

4. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core

outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1–2):9–19. doi:10.1016/j.pain.2004.09.012

5. Crowe M, Whitehead L, Seaton P, Jordan J, Mccall C, Maskill V, Trip H. Qualitative meta-synthesis: the experience of chronic pain across conditions. J Adv Nurs 2017 May;73(5):1004–1016. doi:10.1111/jan.13174

6. Schneider S, Junghaenel DU, Keefe FJ, Schwartz JE, Stone AA, Broderick JE. Individual differences in the day-to-day variability of pain, fatigue, and well-being in patients with rheumatic disease: associations with psychological variables. Pain 2012 Apr;153(4):813–822. doi:10.1016/j.pain.2012.01.001

7. Wesolowicz DM, Bishop MD, Robinson ME. An Examination of Day-to-Day and Intraindividual Pain Variability in Low Back Pain. Pain Med 2021 Oct;22(10):2263–2275. doi:10.1093/pm/pnab119

8. Johnson LM, Zautra AJ, Davis MC. The role of illness uncertainty on coping with fibromyalgia symptoms. Health Psychol 2006 Nov;25(6):696–703. doi:10.1037/0278-6133.25.6.696

9. Little CL, Druce KL, Dixon WG, Schultz DM, House T, McBeth J. What do people living with chronic pain want from a pain forecast? A research prioritisation study. medRxiv 2023 Apr; doi:10.1101/2023.04.24.23289032

10. Flurey CA, Morris M, Richards P, Hughes R, Hewlett S. It's like a juggling act: rheumatoid arthritis patient perspectives on daily life and flare while on current treatment regimes. Rheumatology 2014;53(4):696–703. doi:10.1093/rheumatology/ket416

11. Kongsted A, Kent P, Hestbaek L, Vach W. Patients with low back pain had distinct clinical course patterns that were typically neither complete recovery nor constant pain. A latent class analysis of longitudinal data. Spine J 2015 May;15(5):885–894. doi:10.1016/j.spinee.2015.02.012

12. Dunn KM, Jordan K, Croft PR. Characterizing the Course of Low Back Pain: A Latent Class Analysis. Am J Epidemiol 2006 Apr;163(8):754–761. doi:10.1093/aje/kwj100

13. Bastick AN, Wesseling J, Damen J, Verkleij SPJ, Emans PJ, Bindels PJE, Bierma-Zeinstra SMA. Defining knee pain trajectories in early symptomatic knee osteoarthritis in primary care: 5-year results from a nationwide prospective cohort study (CHECK). Br J Gen Pract 2016 Jan;66(642):e32-39. doi:10.3399/bjgp15X688129

14. Dixon WG, Beukenhorst AL, Yimer BB, Cook L, Gasparrini A, El-Hay T, et al. How the weather affects the pain of citizen scientists using a smartphone app. NPJ Digit Med 2019 Oct;2:105. doi:10.1038/s41746-019-0180-3

15. Check Hayden E. Mobile-phone health apps deliver data bounty. Nature 2016 Mar;531(7595):422–423. doi:10.1038/531422a

16. Shiffman S, Stone AA, Hufford MR. Ecological Momentary Assessment. Annu Rev Clin Psychol 2008;4:1–32. doi:10.1146/annurev.clinpsy.3.022806.091415

17. Liddell TM, Kruschke JK. Analyzing ordinal data with metric models: What could possibly go

wrong? J Exp Soc Psychol 2018;79:328–348. doi:10.1016/j.jesp.2018.08.009

18. Schultz DM, Beukenhorst AL, Yimer BB, Cook L, Pisaniello HL, House T, et al. Weather Patterns Associated with Pain in Chronic-Pain Sufferers. Bull Am Meteorol Soc 2020;101(5):E555–E566. doi:10.1175/BAMS-D-19-0265.1

19. Druce KL, McBeth J, van der Veer SN, Selby DA, Vidgen B, Georgatzis K, et al. Recruitment and Ongoing Engagement in a UK Smartphone Study Examining the Association Between Weather and Pain: Cohort Study. JMIR MHealth UHealth 2017 Nov;5(11):e168. doi:10.2196/mhealth.8162

20. Knecht C, Hartnack S, Sick B, Riner F, Schweinhardt P, Wirth B. A prospective observational study on trajectories and prognostic factors of mid back pain. BMC Musculoskelet Disord 2020 Aug;21(1):554. doi:10.1186/s12891-020-03534-5

21. Crook J, Moldofsky H. The clinical course of musculoskeletal pain in empirically derived groupings of injured workers. Pain 1996 Oct;67(2–3):427–433. doi:10.1016/0304-3959(96)03151-x

22. Axén I, Bodin L, Bergström G, Halasz L, Lange F, Lövgren PW, et al. Clustering patients on the basis of their individual course of low back pain over a six month period. BMC Musculoskelet Disord 2011 May;12:99. doi: 10.1186/1471-2474-12-99

23. Brennan PL, SooHoo S. Effects of Mental Health Disorders on Nursing Home Residents' Nine-Month Pain Trajectories. Pain Med 2020 Mar;21(3):488–500. doi:10.1093/pm/pnz177

24. Ogollah RO, Konstantinou K, Stynes S, Dunn KM. Determining One-Year Trajectories of Low-Back-Related Leg Pain in Primary Care Patients: Growth Mixture Modeling of a Prospective Cohort Study. Arthritis Care Res 2018 Dec;70(12):1840–1848. doi:10.1002/acr.23556

25. Battaglia M, Garon-Carrier G, Brendgen M, Feng B, Dionne G, Vitaro F, et al. Trajectories of pain and anxiety in a longitudinal cohort of adolescent twins. Depress Anxiety 2020 May;37(5):475–484. doi:10.1002/da.22992

26. Neupane S, Nygård C-H, Prakash KC, von Bonsdorff MB, von Bonsdorff ME, Seitsamo J, et al. Multisite musculoskeletal pain trajectories from midlife to old age: a 28-year follow-up of municipal employees. Occup Environ Med 2018 Dec;75(12):863–870. doi:10.1136/oemed-2018-105235

27. Dunn KM, Campbell P, Jordan KP. Long-term trajectories of back pain: cohort study with 7-year follow-up. BMJ Open 2013 Dec;3(12):e003838. doi:10.1136/bmjopen-2013-003838

28. Hallman DM, Rasmussen CDN, Jørgensen MB, Holtermann A. Time course of neck-shoulder pain among workers: A longitudinal latent class growth analysis. Scand J Work Environ Health 2018 Jan;44(1):47–57. doi:10.5271/sjweh.3690

29. Hebert JJ, Abraham E, Wedderkopp N, Bigney E, Richardson E, Darling M, et al. Patients undergoing surgery for lumbar spinal stenosis experience unique courses of pain and disability: A group-based trajectory analysis. PloS One 2019 Nov;14(11):e0224200. doi:10.1371/journal.pone.0224200

30. Wieczorek M, Rotonda C, Coste J, Pouchot J, Saraux A, Guillemin F, Rat A-C.Trajectory analysis combining pain and physical function in individuals with knee and hip osteoarthritis: results from the French KHOALA cohort. Rheumatol 2020 Nov;59(11):3488–3498. doi:10.1093/rheumatology/keaa148

31. Dai Z, Lu N, Niu J, Felson DT, Zhang Y. Dietary Fiber Intake in Relation to Knee Pain Trajectory. Arthritis Care Res 2017 Sep;69(9):1331–1339. doi:10.1002/acr.23158

32. Radojčić MR, Arden NK, Yang X, Strauss VY, Birrell F, Cooper C, et al. Pain trajectory defines knee osteoarthritis subgroups: a prospective observational study. Pain 2020 Dec;161(12):2841–2851. doi:10.1097/j.pain.0000000000001975

33. Pan F, Tian J, Aitken D, Cicuttini F, Jones G. Predictors of pain severity trajectory in older adults: a 10.7-year follow-up study. Osteoarthritis Cartilage 2018 Dec;26(12):1619–1626. doi:10.1016/j.joca.2018.08.002

34. Pan F, Tian J, Cicuttini F, Jones G. Metabolic syndrome and trajectory of knee pain in older adults. Osteoarthritis Cartilage 2020 Jan;28(1):45–52. doi:10.1016/j.joca.2019.05.030

35. Kassambara A. Practical guide to cluster analysis in R: unsupervised machine learning, Edition 1. United States: STHDA; 2017. ISBN:9781542462709

36. Kaufman L, Rousseeuw PJ. Clustering Large Data Sets. Pattern Recognition in Practice II: proceedings of an international workshop held in Amsterdam, June 19–21, 1985. Elsevier; 1986. p. 425–437. ISBN:9780444878779

37. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. Vol.344. Hoboken, N.J: Wiley; 2005. ISBN:9780471735786

38. Sarang P. Centroid-Based Clustering: Clustering Algorithms for Hard Clustering. In: Thinking Data Science. Cham: Springer International Publishing; 2023. p. 171–183. ISBN:9783031023637

39. Hardy A. On the number of clusters. Comput Stat Data Anal 1996 Nov;23(1):83–96. doi:10.1016/S0167-9473(96)00022-9

40. Thomas MJ, Yu D, Nicholls E, Bierma-Zeinstra S, Conaghan PG, Stoner KJ, et al. Short-Term Recovery Trajectories of Acute Flares in Knee Pain: A UK-Netherlands Multicenter Prospective Cohort Analysis. Arthritis Care Res 2020 Dec;72(12):1687–1692. doi:10.1002/acr.24088

41. Airila A, Hakanen JJ, Luukkonen R, Lusa S, Punakallio A, Leino-Arjas P. Developmental trajectories of multisite musculoskeletal pain and depressive symptoms: the effects of job demands and resources and individual factors. Psychol Health 2014;29(12):1421–1441. doi:10.1080/08870446.2014.945929

42. James RJE, Walsh DA, Ferguson E. General and disease-specific pain trajectories as predictors of social and political outcomes in arthritis and cancer. BMC Med 2018 Apr;16(1):51. doi:10.1186/s12916-018-1031-9

43. Canizares M, Rampersaud YR, Badley EM. Course of Back Pain in the Canadian Population: Trajectories, Predictors, and Outcomes. Arthritis Care Res 2019 Dec;71(12):1660–1670.

doi:10.1002/acr.23811

44. Genolini C, Falissard B. Kml: A package to cluster longitudinal data. Comput Methods Programs Biomed 2011 Dec;104(3):e112–e121. doi:10.1016/j.cmpb.2011.05.008

45. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull World Health Organ 2007 Nov;85(11):867–872. doi:10.2471/blt.07.045120

46. Verkleij SPJ, Hoekstra T, Rozendaal RM, Waarsing JH, Koes BW, Luijsterburg PAJ, Bierma-Zeinstra SMA. Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. Ann Rheum Dis 2012 Sep;71(9):1517–1523. doi:10.1136/annrheumdis-2011-200687

47. Wesseling J, Bastick AN, ten Wolde S, Kloppenburg M, Lafeber FPJG, Bierma-Zeinstra SMA, Bijlsma JWJ. Identifying Trajectories of Pain Severity in Early Symptomatic Knee Osteoarthritis: A 5-year Followup of the Cohort Hip and Cohort Knee (CHECK) Study. J Rheumatol 2015 Aug;42(8):1470–1477. doi:10.3899/jrheum.141036

48. Nicholls E, Thomas E, van der Windt DA, Croft PR, Peat G. Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthritis Cartilage 2014 Dec;22(12):2041–2050. doi:10.1016/j.joca.2014.09.026

49. Bastick AN, Verkleij SPJ, Damen J, Wesseling J, Hilberdink WKHA, Bindels PJE, Bierma-Zeinstra SMA. Defining hip pain trajectories in early symptomatic hip osteoarthritis – 5 year results from a nationwide prospective cohort study (CHECK). Osteoarthritis Cartilage 2016 May;24(5):768–775. doi:10.1016/j.joca.2015.11.023

50. Collins JE, Katz JN, Dervan EE, Losina E. Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2014 May;22(5):622–630. doi:10.1016/j.joca.2014.03.009

51. Johnson AJ, Vasilopoulos T, Booker SQ, Cardoso J, Terry EL, Powell-Roach K, et al. Knee pain trajectories over 18 months in non-Hispanic Black and non-Hispanic White adults with or at risk for knee osteoarthritis. BMC Musculoskelet Disord 2021 May;22(1):415. doi:10.1186/s12891-021-04284-8

52. Atukorala I, Downie A, Pathmeswaran A, Deveza LMA, Chang T, Zhang Y, Hunter DJ. Short-term pain trajectories in patients with knee osteoarthritis. Int J Rheum Dis 2022 Mar;25(3):281–294. doi:10.1111/1756-185X.14267

53. Tamcan O, Mannion AF, Eisenring C, Horisberger B, Elfering A, Müller U. The course of chronic and recurrent low back pain in the general population. Pain 2010 Sep;150(3):451–457. doi:10.1016/j.pain.2010.05.019

54. Coenen P, Smith A, Paananen M, O'Sullivan P, Beales D, Straker L. Trajectories of Low Back Pain From Adolescence to Young Adulthood. Arthritis Care Res 2017 Mar;69(3):403–412. doi:10.1002/acr.22949

55. Ailliet L, Rubinstein SM, Hoekstra T, van Tulder MW, de Vet HCW. Long-term trajectories of

patients with neck pain and low back pain presenting to chiropractic care: A latent class growth analysis. Eur J Pain 2018 Jan;22(1):103–113. doi: 10.1002/ejp.1094

56. Downie AS, Hancock MJ, Rzewuska M, Williams CM, Lin C-WC, Maher CG. Trajectories of acute low back pain: a latent class growth analysis. Pain 2016 Jan;157(1):225–234. doi:10.1097/j.pain.0000000000000351

57. Panken G, Hoekstra T, Verhagen A, van Tulder M, Twisk J, Heymans MW. Predicting chronic low-back pain based on pain trajectories in patients in an occupational setting: an exploratory analysis. Scand J Work Environ Health 2016 Jun;42(6):520–527. doi:10.5271/sjweh.3584

58. Macedo LG, Maher CG, Latimer J, McAuley JH, Hodges PW, Rogers WT. Nature and determinants of the course of chronic low back pain over a 12-month period: a cluster analysis. Phys Ther 2014 Feb;94(2):210–221. doi:10.2522/ptj.20120416

59. Kyrönlahti SM, Nygård C-H, Prakash K C, Neupane S. Trajectories of low back pain from midlife to retirement and functional ability at old age. Eur J Public Health 2022 Jun;32(3):497–503. doi:10.1093/eurpub/ckab191

60. Enthoven WTM, Koes BW, Bierma-Zeinstra SMA, Bueving HJ, Bohnen AM, Peul WC, et al. Defining trajectories in older adults with back pain presenting in general practice. Age Ageing 2016 Nov;45(6):878–883. doi:10.1093/ageing/afw127

61. Pico-Espinosa OJ, Côté P, Hogg-Johnson S, Jensen I, Axén I, Holm LW, Skillgate E. Trajectories of Pain Intensity Over 1 Year in Adults With Disabling Subacute or Chronic Neck Pain. Clin J Pain 2019 Aug;35(8):678–685. doi:10.1097/AJP.0000000000000727

62. Eriks-Hoogland IE, Hoekstra T, de Groot S, Stucki G, Post MW, van der Woude LH. Trajectories of musculoskeletal shoulder pain after spinal cord injury: Identification and predictors. J Spinal Cord Med 2014 May;37(3):288–298. doi:10.1179/2045772313Y.0000000168

63. Haybatollahi SM, James RJE, Fernandes G, Valdes A, Doherty M, Zhang W, et al. Identifying multiple knee pain trajectories and the prediction of opioid and NSAID medication used: A latent class growth approach. Pain Pract 2022 Feb;22(2):210–221. doi:10.1111/papr.13082

64. Marshall M, Blagojevic-Bucknall M, Rathod-Mistry T, Thomas MJ, Edwards JJ, Peat G, et al. Identifying Long-Term Trajectories of Foot Pain Severity and Potential Prognostic Factors: A Population-Based Cohort Study. Arthritis Care Res 2023 May;75(5). doi:10.1002/acr.24823

65. Leino-Arjas P, Rajaleid K, Mekuria G, Nummi T, Virtanen P, Hammarström A. Trajectories of musculoskeletal pain from adolescence to middle age: the role of early depressive symptoms, a 27-year follow-up of the Northern Swedish Cohort. Pain 2018 Jan;159(1):67–74. doi:10.1097/j.pain.0000000000001065

66. Oakman J, Neupane S, Kyrönlahti S, Nygård C-H, Lambert K. Musculoskeletal pain trajectories of employees working from home during the COVID-19 pandemic. Int Arch Occup Environ Health 2022 Nov;95(9):1891–1901. doi:10.1007/s00420-022-01885-1

67. Neupane S, Lallukka T, Pietiläinen O, Rahkonen O, Leino-Arjas P. Trajectories of multisite musculoskeletal pain in midlife: Associations with common mental disorders. Eur J Pain 2020

Feb;24(2):364–373. doi:10.1002/ejp.1492

68. Sun F, Park NS, Wardian J, Lee BS, Roff LL, Klemmack DL, et al. Predicting the Trajectories of Perceived Pain Intensity in Southern Community-Dwelling Older Adults: The Role of Religiousness. Res Aging 2013 Nov;35(6). doi:10.1177/0164027512456402

69. Wickrama KAS, Lee TK, O'Neal CW. Heterogeneity in Husbands' and Wives' Physical Pain Trajectories Over Mid-Later Years: Biopsychosocial Stratification and Implications for Later-Life Well-Being. The Gerontologist 2021 Aug;61(6):930–941. doi:10.1093/geront/gnab007

70. Kongsted A, Hestbæk L, Kent P. How can latent trajectories of back pain be translated into defined subgroups? BMC Musculoskelet Disord 2017 Jul;18(1):285. doi:10.1186/s12891-017-1644-8

71. Beukenhorst AL, Schultz DM, McBeth J, Sergeant JC, Dixon WG. Are weather conditions associated with chronic musculoskeletal pain? Review of results and methodologies. Pain 2020 Apr;161(4):668–683. doi:10.1097/j.pain.0000000000001776

72. Genolini C, Alacoque X, Sentenac M, Arnaud C. kml and kml3d: *R* Packages to Cluster Longitudinal Data. J Stat Softw 2015;65(4):1–34. doi:10.18637/jss.v065.i04

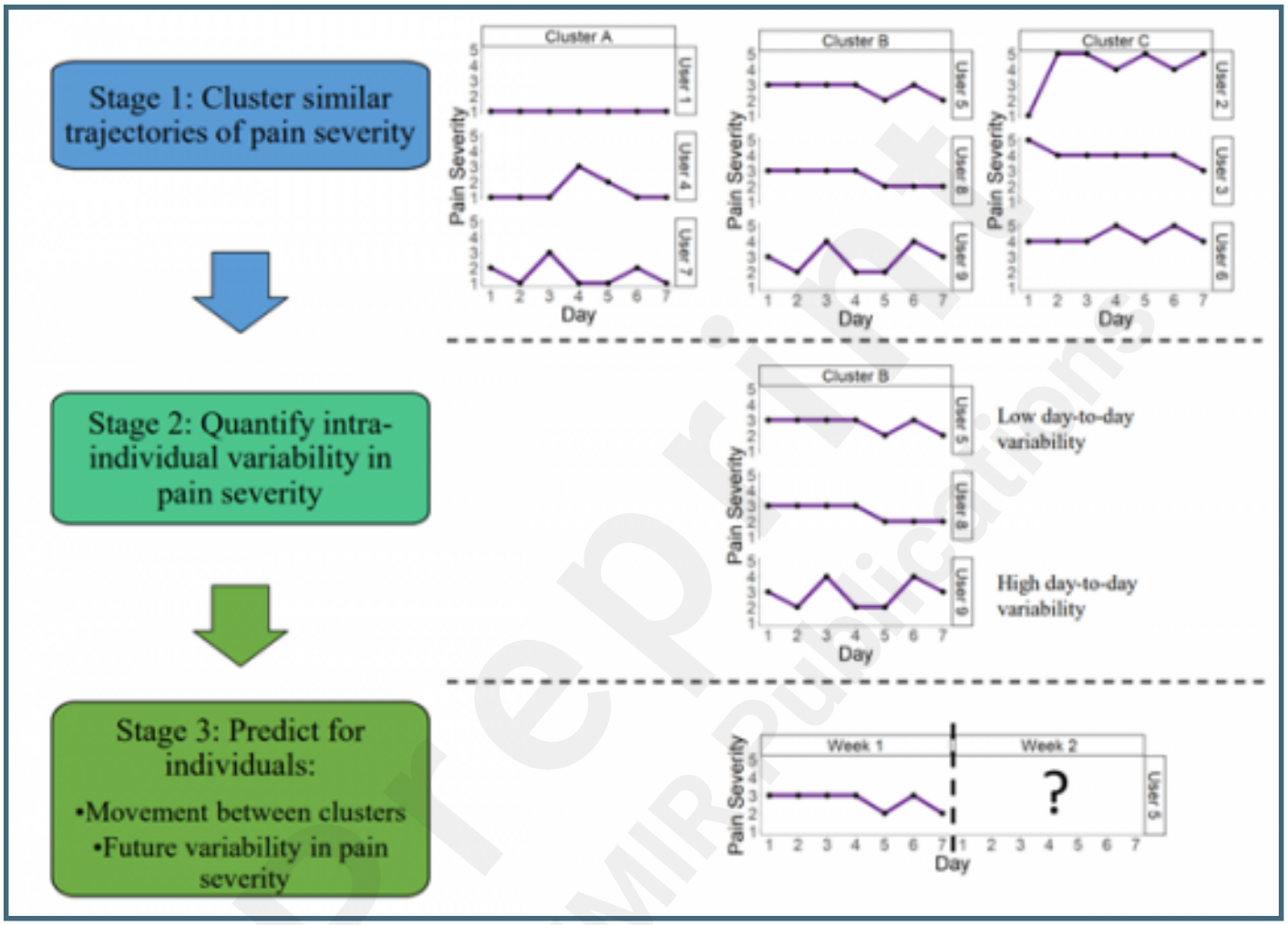
Abbreviations
App: application
mHealth: mobile health
s.d.: standard deviation
WSS: within-cluster sum of squares

Supplementary Files

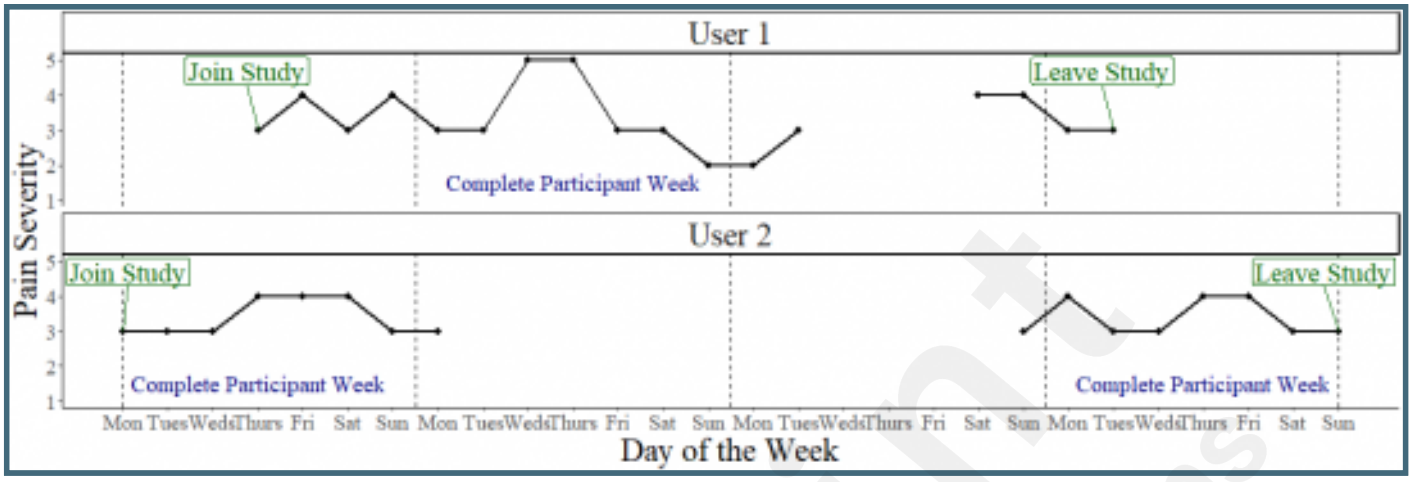
Revised manuscript with tracked changes as of 12.11.23.
URL: http://asset.jmir.pub/assets/166e7dde119deb565d6b242f1fac796f.docx

Figures

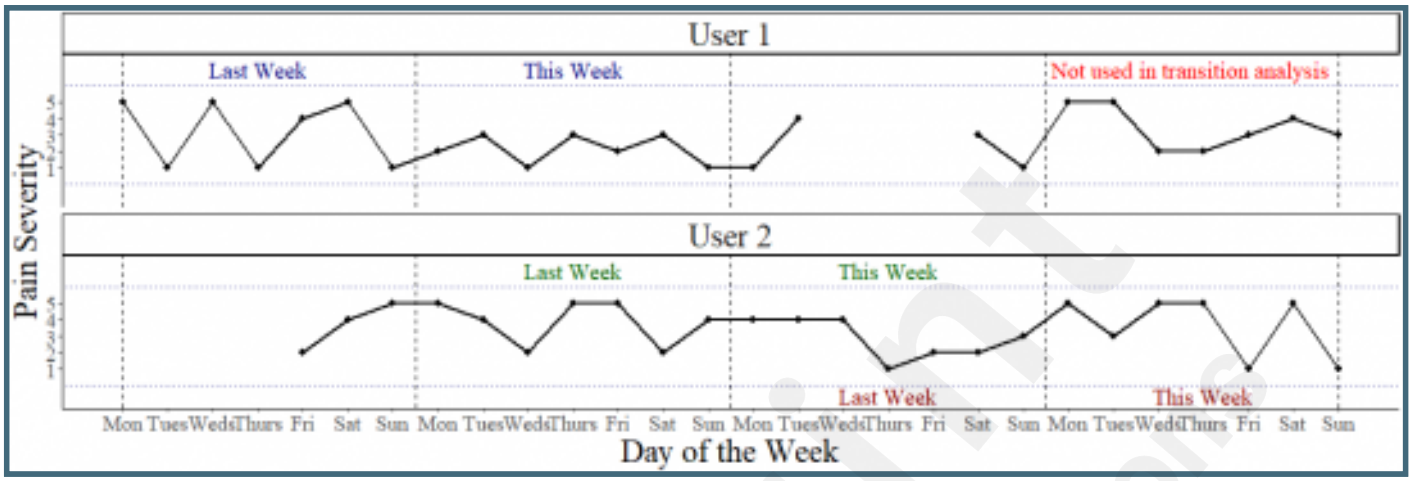
Three stages to build a pain forecast using data from an mHealth study. Data used in this figure are for illustrative purposes only, to provide one example of how data may be used in the pipeline of developing a pain forecast. First, data are clustered to identify common trajectories of weekly pain severity. Second, the remaining variability is explored for each trajectory within a cluster. The process is repeated for each cluster. Third, movement between clusters on consecutive weeks and the amount of day-to-day variability are predicted for an individual. The process is repeated for each individual.



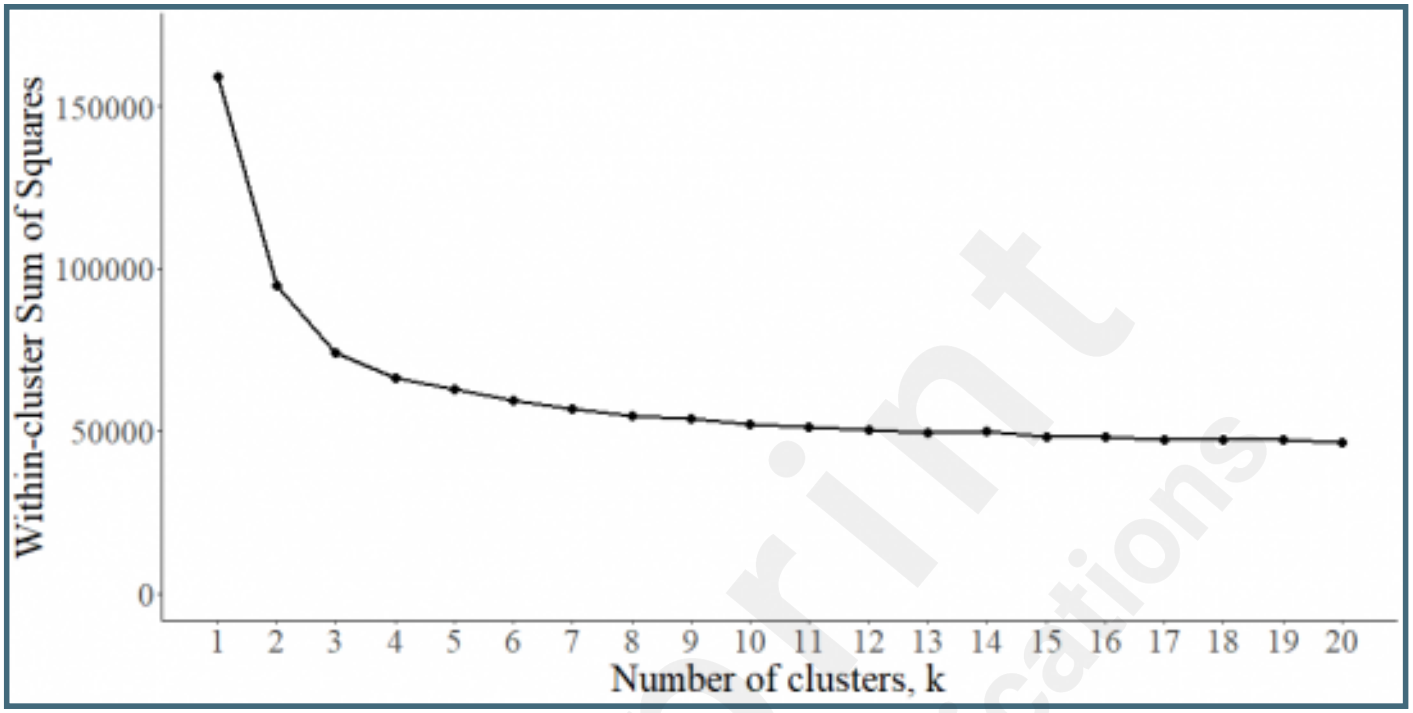
Example selections of complete trajectory-weeks for two participants. The participants join and leave the study at different times. One complete participant-week from User 1 is included in the analysis. Two complete participant-weeks from User 2 are included in the analysis.



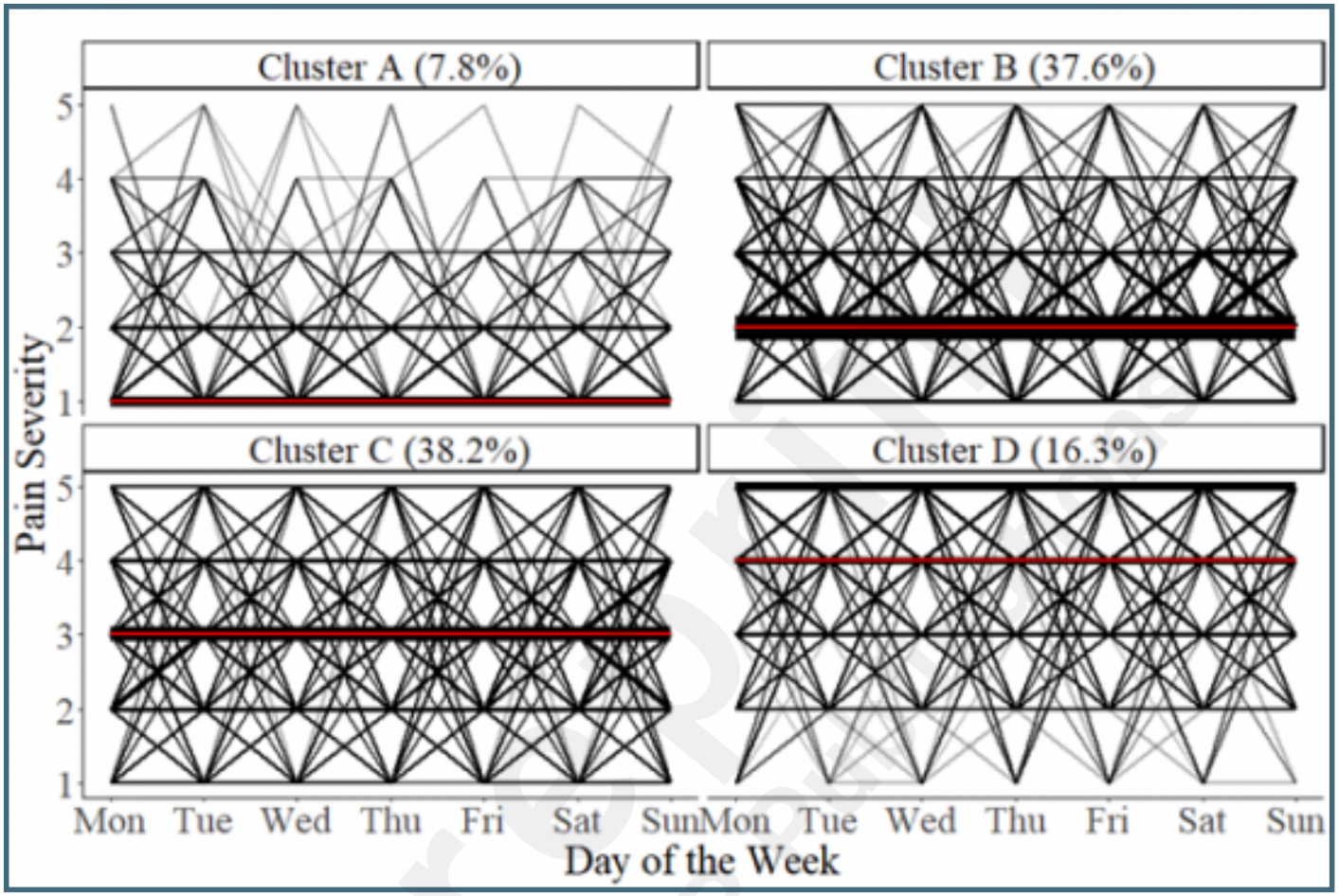
Example data from two participants, highlighting how their data is used to examine transitions between clusters. User 1 provided data in three complete participant-weeks. The first two are consecutive and therefore are used in the transition analysis. The final complete participant-week is not used. User 2 provided three complete participant-weeks. All three are used in the transition analysis. The middle week is labelled as both ‘This Week’ and ‘Last Week’ in different pairings.



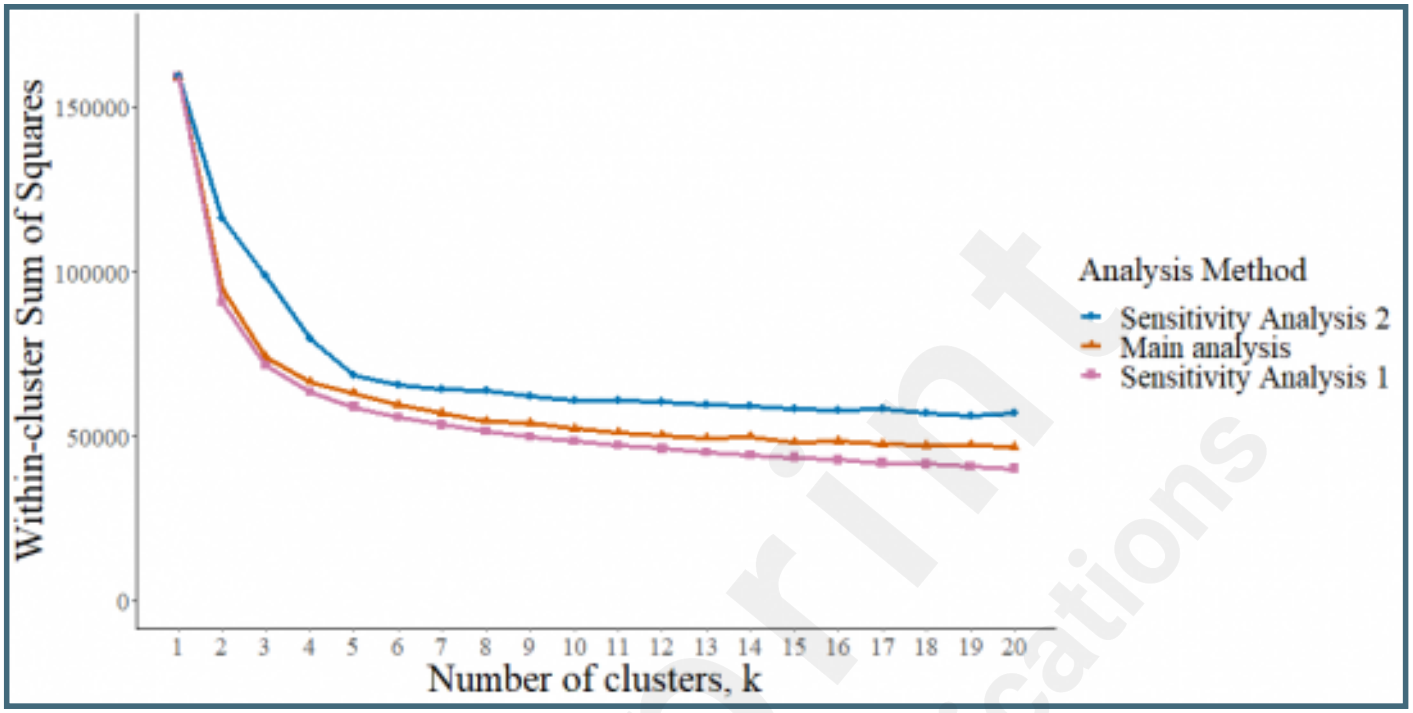
Unexplained variability across different cluster (k) solutions. The within-sum-of-squares indicates the remaining variance within clusters. An elbow at k = 4 suggests an appropriate solution, with diminishing returns for the inclusion of further clusters.



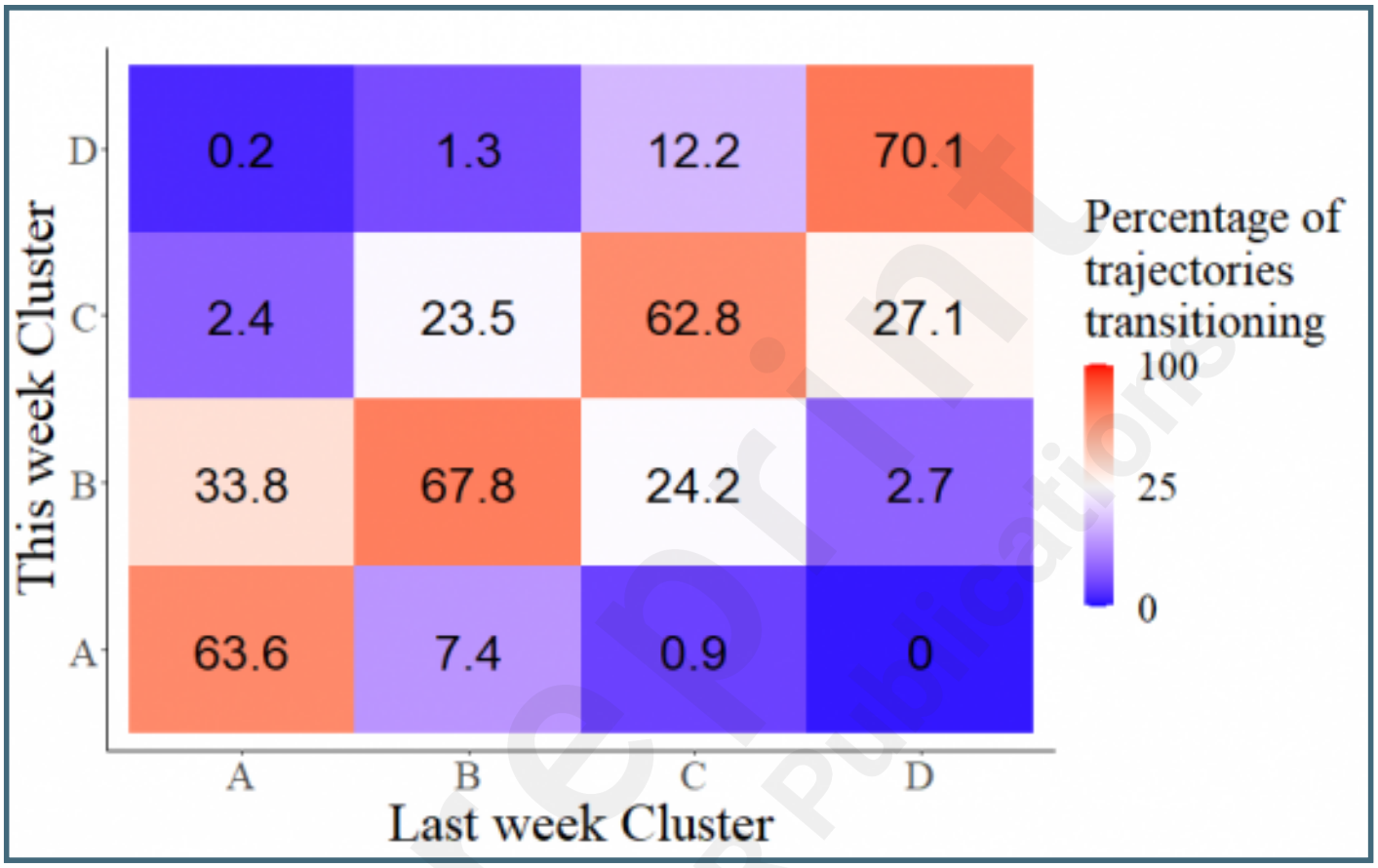
Weighted spaghetti plot of trajectories assigned to each cluster. The weight (and transparency) of each path represents the number of trajectories following that path. The red line represents the medoid of the cluster. Cluster A = “no or low pain”, B = “mild pain”, C = “moderate pain” and D = “severe pain”. The percentage of trajectories assigned to each cluster is shown.



Unexplained variability across different cluster (k) solutions for the main analysis and two sensitivity analyses. In the main analysis and Sensitivity Analysis 1, there is an elbow at k = 4. In Sensitivity Analysis 2, there is an elbow at k = 5. Separate graphs are provided in Supplementary Materials 2.



Transition matrix of movement between clusters on consecutive weeks. For membership in each cluster “Last Week”, the percentage of membership in each cluster “This Week” is shown. Random movement between clusters would suggest each combination has a transition percentage of 25%. Blue squares represent transitions that have a higher than expected percentage (> 25%). Red squares represent transitions that have a lower than expected percentage (< 25%). A white square would have exactly the expected percentage (25%).



Multimedia Appendixes

Untitled.
URL: http://asset.jmir.pub/assets/c0ba2c358c562e58d9fad0e778b803a2.docx

Supplementary materials.
URL: http://asset.jmir.pub/assets/a4f5afac6848acf220cc6e5c87bb865d.docx